On the Lifting and Approximation Theorem for Nonsmooth Vector Fields

MARCO BRAMANTI, LUCA BRANDOLINI & MARCO PEDRONI

ABSTRACT. We prove a version of Rothschild-Stein's theorem of lifting and approximation and some related results in the context of nonsmooth Hörmander's vector fields for which the highest order commutators are only Hölder continuous. The theory explicitly covers the case of one vector field having weight two while the others have weight one.

INTRODUCTION

This paper is focussed on the well-known "lifting and approximation" theorem proved by Rothschild-Stein in [14], and related topics. To describe the context and aim of this paper, we have therefore to recall what that theorem is about. (Here we will be rather sketchy, while precise definitions will be given later).

Let us consider a family of real smooth vector fields $X_0, X_1, \dots, X_n$ defined in some domain of $\mathbb{R}^p$, and the corresponding second order differential operator

$$L = \sum_{i=1}^{n} X_i^2 + X_0. \tag{0.1}$$

If the X_i's satisfy Hörmander's condition, then L is hypoellitptic (Hörmander's theorem, [10]).

If there exists in $\mathbb{R}^p$ a structure of "homogeneous group" such that L is left invariant (with respect to the group translations) and homogeneous of degree two (with respect to the group dilations), then L possesses a homogeneous fundamental solution (Folland, [7]) which allows one to apply fairly standard techniques of singular integrals, in order to prove a-priori estimates and other interesting properties of L. If such a group structure does not exists, then Rothschild-Stein's theory tells us that it is still possible to reduce, in a suitable sense, the study of L to the study of a homogeneous left invariant operator. This requires a three-step process. First, one "lifts" the original vector fields

$$X_i = \sum_{j=1}^{p} a_{ij}(x)\,\partial_{x_j}, \quad x \in \mathbb{R}^p$$

which are assumed to satisfy Hörmander's condition at some step r, to some new vector fields

$$\tilde{X}_i = X_i + \sum_{j=1}^{m} b_{ij}(x,t)\,\partial_{t_j}, \quad (x,t) \in \mathbb{R}^{p+m}$$

so that these lifted vector fields still satisfy Hörmander's condition at the same step r, and are free up to step r.

Second, one proves that in $\mathbb{R}^{p+m}$ there exists a structure of homogeneous group G and a family of left invariant homogeneous vector fields Y_i which locally approximate the $\tilde{X}_i$'s. More precisely, for every point $\eta = (x,t)$ there is a local diffeomorphism

$$u = \Theta_\eta(\xi)$$

from a neighborhood of η onto a neighborhood of the origin in G, such that with respect to these local coordinates,

$$\tilde{X}_i = Y_i + R_i^\eta$$

where the "remainder" R_i^η is a vector field, smoothly depending on the parameter η, such that its action on the fundamental solution Γ of

$$\sum_{i=1}^{n} Y_i^2 + Y_0$$

gives a function which is less singular than $Y_i\Gamma$. The map $\Theta_\eta(\cdot)$, a key object in this theory, also possesses other interesting properties:

(1) It depends smoothly on η;

(2) The function

$$\rho(\xi,\eta) = \|\Theta_\eta(\xi)\|$$

(where $\|\cdot\|$ is a homogeneous norm on the group G) is a quasidistance (which also turns out to be equivalent to the distance induced by the vector fields);

(3) The change of variables $u = \Theta_\eta(\xi)$ obeys to

$$d\xi = c(\eta)(1 + O(\|u\|))\, du$$

where the function $c(\eta)$ is smooth and bounded away from zero.

The set of results just described allows one to prove suitable a priori estimates for the lifted operator $\tilde{L}$. Once these are proved, it is not difficult (third step) to derive the corresponding estimates for the original operator, exploiting the fact that L is the projection of $\tilde{L}$ on $\mathbb{R}^p$. All these results have been proved in [14].

Over the years, the lifting and approximation technique has showed to be useful also for other purposes. In particular, sometimes the lifting theorem is enough (without need of the approximation part of the theory), in order to reduce problems for a general family of Hörmander's vector fields to problems for free vector fields, which for algebraic reasons are easier to be studied.

Since the original proof of the lifting and approximation theorem given in [14] is long and difficult, several authors have given alternative proofs: Hörmander-Melin [11], Folland [8] and Goodman [9] prove the lifting theorem and a pointwise version of the approximation theorem, without dealing with the map $\Theta_\eta(\cdot)$ and its properties; Folland restricts to the particular case when the starting vector fields are already left invariant and homogeneous with respect to a group structure (but are not free); more recently, Bonfiglioli-Uguzzoni [4] have proved that under Folland's assumptions, the original vector fields can be lifted directly to free left invariant homogeneous vector fields (in other words, in this case the "remainders" R_i^η can be taken equal to zero). Coming back to the case of general Hörmander's vector fields, Christ-Nagel-Stein-Weinger [5] prove a somewhat more general version of the lifting theorem, because they consider "weighted" vector fields (we will explain this feature in a moment); on the other hand, they do not prove any approximation result.

Although Rothschild-Stein state their main results for a Hörmander operator (0.1), all their proofs are written for the "sum of squares" operator

$$L = \sum_{i=1}^{n} X_i^2.$$

The issue in handling Hörmander's operators (0.1) consists in the fact that the vector field X_0 has "weight" 2, while $X_1, X_2, \ldots, X_n$ have weight 1; this fact requires to modify in a suitable way all the basic definitions appearing in this context (free vector fields, weight of a commutator,...); due to the complexity of the theory, this adaptation is not trivial. Nevertheless, as far as we know, a detailed proof of lifting, approximation, and properties of the map $\Theta_\eta(\cdot)$, adapted to the case of weighted vector fields has not been written yet (as we have already pointed out, the paper [5] considers weighted vector fields but only contains a proof of the lifting result).

A first aim of this paper is to present a detailed proof of the aforementioned results, explicitly covering the case of *weighted vector fields*.

Second, we are interested in extending these results to the case of *nonsmooth* vector fields, that is, vector fields which only possess the number of derivatives involved in the commutators which are necessary to check Hörmander's condition, with Hölder continuous derivatives of the maximum order. This is part of a larger project which we have started in [2], where we have proved in this nonsmooth context a Poincaré's inequality, together with the basic properties of the distance induced by the X_i's: Chow's connectivity theorem, the doubling condition, the equivalence between different distances induced by the X_i's, etc. We refer to the introduction of [2] for a survey of the existing literature about nonsmooth Hörmander's vector fields.

We point out that the *lifting* theorem which we prove here has already been used in [2], as one of the tools in the proof of a Poincaré's inequality. Moreover, the whole set of results proved in this paper allows to establish, for the operator L, the existence of a local fundamental solution. This result, obtained by a suitable adaptation of the Levi's parametrix method, will be accomplished in the forthcoming paper [3].

The main results about nonsmooth weighted vector fields proved in this paper are: lifting (Theorem 1.5, Section 1.2), approximation (Theorem 3.9, Section 3.3), and properties of the map $\Theta_\eta(\cdot)$ (Proposition 3.8, Section 3.3, Proposition 3.10 and Proposition 3.11, Section 3.4), which are the analog of the properties (1), (2), (3) quoted at the beginning of the Introduction.

We now describe the general strategy that we have followed and the structure of the paper.

In Section 1 we prove the lifting theorem for weighted nonsmooth Hörmander's vector field. Here and in the following section we adapt and detail the arguments in [11]. However, in order to prove the approximation result for nonsmooth vector fields, it is not possible to proceed further in the line of [11]. The reason is that a basic idea of this theory is that of rewriting the vector fields in "canonical coordinates"; this means to apply a suitable change of variables, which however turns out to be just Hölder continuous if the vector fields have the limited smoothness that we assume, so that this way is closed.

Therefore, in Section 2, we pass to consider free smooth vector fields, proving for them the approximation result, a ball-box theorem, and the desired properties of the map $\Theta_\eta(\cdot)$. As a by-product we also get, in the particular case of free vector fields, a quite simple proof of the results due to Nagel-Stein-Wainger [12] about the volume of metric balls and the doubling condition.

Then, in Section 3, we come back to nonsmooth free vector fields. Now the natural idea is to approximate nonsmooth vector fields with smooth ones, obtained taking suitable Taylor's expansions of the coefficients. This idea, firstly introduced in [6], has been used also in [2]. To these approximating smooth vector fields we can apply the theory developed in Section 2, in order to derive the corresponding results in the nonsmooth case. We stress the fact that, in the nonsmooth context, the properties of the map $\Theta_\eta(\cdot)$ hold in a weaker form: the dependence on η is only Hölder continuous. However, this is enough to get the aforementioned existence result for a fundamental solution for L.

1. Lifting of Nonsmooth Hörmander's Vector Fields

1.1. Assumptions and notation. Let $X_0, X_1, \dots, X_n$ be a system of real vector fields, defined in a domain of $\mathbb{R}^p$. Let us assign to each X_i a *weight* p_i, saying that

$$p_0 = 2 \text{ and } p_i = 1 \quad \text{for } i = 1, 2, \dots, n.$$

The following standard notation, will be used throughout the paper. For any multiindex

$$I = (i_1, i_2, \dots, i_k)$$

we define the *weight* of I as

$$|I| = \sum_{j=1}^{k} p_{i_j}.$$

Sometimes, we will also use the (usual) *length* of I,

$$\ell(I) = k.$$

For any vector field X, we denote by $\operatorname{ad} X$ the linear operator which maps Y to $[X, Y]$, where Y is any vector field and $[\cdot, \cdot]$ is the Lie bracket. Now, for any multiindex $I = (i_1, i_2, \dots, i_k)$ we set:

$$X_I = X_{i_1} X_{i_2} \dots, X_{i_k}$$

and

$$X_{[I]} = \operatorname{ad} X_{i_1} \operatorname{ad} X_{i_2} \dots \operatorname{ad} X_{i_{k-1}} X_{i_k} = [X_{i_1}, [X_{i_2}, \dots [X_{i_{k-1}}, X_{i_k}] \dots]].$$

If $I = (i_1)$, then

$$X_{[I]} = X_{i_1} = X_I.$$

As usual, $X_{[I]}$ can be seen either as a differential operator or as a vector field. We will write

$$X_{[I]} f$$

to denote the differential operator $X_{[I]}$ acting on a function f, and

$$(X_{[I]})_x$$

to denote the vector field $X_{[I]}$ evaluated at the point x.

Assumptions (A). We assume that for some integer $r \geq 2$ and some bounded domain (i.e., connected open subset) $\Omega \subset \mathbb{R}^p$ the following hold:

The coefficients of the vector fields $X_1, X_2, \dots, X_n$ belong to $C^{r-1}(\Omega)$, while the coefficients of X_0 belong to $C^{r-2}(\Omega)$. Here and in the following, C^k stands for the classical space of functions with continuous derivatives up to order k.

These assumptions are consistent in view of the following result.

Lemma 1.1. *Under the assumption (A) above, for any $1 \leq k \leq r$, the differential operators*

$$\{X_I\}_{|I| \leq k}$$

are well defined, and have C^{r-k} coefficients. The same is true for the vector fields $\{X_{[I]}\}_{|I| \leq k}$.

Proof. By induction on k. For $k = 1$, the assertion is part of the assumption (A). Assume the assertion holds up to $k - 1$, and let

$$I = (i_1, i_2, \dots, i_m), \quad \text{with } |I| = k.$$

Set $I' = (i_2, \dots, i_m)$ so that $X_I(x) = X_{i_1} X_{I'}(x)$. If X_{i_1} has weight p_{i_1}, then $|I'| = k - p_{i_1} \geq 0$; by inductive assumption, $X_{I'}$ has $C^{r-k+p_{i_1}}$ coefficients, hence $X_{i_1} X_{I'}(x)$ has C^h coefficients, with $h = \min(r - k + p_{i_1} - 1, r - p_{i_1}) \geq r - k$, and we are done. □

1.2. Hörmander-Melin procedure. We are now going to define the concept of free vector fields. Clearly, any vector field $X_{[I]}$ with $|I| \leq r$ can be rewritten explicitly as a linear combination of operators of the kind X_J for $|J| = |I|$:

$$X_{[I]} = \sum_J A_{IJ} X_J$$

where $\{A_{IJ}\}_{|I|,|J| \leq r}$ is a matrix of universal constants, built exploiting only those relations between $X_{[I]}$ and X_J which hold automatically, as a consequence of the definition of $X_{[I]}$, regardless of the specific properties of the vector fields $X_0, X_1, \dots, X_n$. In particular, we see that

$$A_{IJ} = 0 \qquad \text{if } |J| \neq |I|$$

and

$$A_{IJ} = \delta_{IJ} \quad \text{if } |J| = |I| = 1.$$

Also, note that if $\{a_I\}_{I \in B}$ is any finite set of constants such that

$$\sum_{I \in B} a_I A_{IJ} = 0 \ \forall J, \quad \text{then } \sum_{I \in B} a_I X_{[I]} \equiv 0 \tag{1.1}$$

for arbitrary vector fields $X_0, X_1, \dots, X_n$. Reversing this property we get the definition of a key concept which will be dealt with in the following:

Definition 1.2. For any positive integer $s \le r$, we say that the vector fields $X_0, X_1, \dots, X_n$ are free up to weight s at 0, if, for any family of constants $\{a_I\}_{|I|\le s}$,

$$\sum_{|I|\le s} a_I (X_{[I]})_0 = 0 \;\Rightarrow\; \sum_{|I|\le s} a_I A_{IJ} = 0 \quad \forall\, J.$$

Comparing this definition with (1.1) shows that $X_0, X_1, \dots, X_n$ are free up to weight s at 0 if, roughly speaking, the only linear identities relating the $X_{[I]}$'s for $|I| \le s$ (at 0) are those which hold for any possible choice of $X_0, X_1, \dots, X_n$, as a consequence of the formal properties of the Lie bracket, namely antisymmetry and Jacobi's identity. However, the different weights of the vector fields make this property not so easy to state more explicitly. For instance, saying that $X_0, X_1, \dots, X_n$ are free up to weight 1 at 0 just means that $X_1, \dots, X_n$ are linearly independent (without any requirement on X_0); saying that $X_0, X_1, \dots, X_n$ are free up to weight 2 at 0 means that:

(i) $X_1, \dots, X_n$ are linearly independent;
(ii) X_0 is not a linear combination of vector fields of the kind $[X_i, X_j]$ for $i, j = 1, 2, \dots, n$;
(iii) The only linear relations between the $[X_i, X_j]$'s (for $i, j = 1, 2, \dots, n$) are those following from antisymmetry.

Clearly, for a general s the explicit description of this property becomes cumbersome.

Proposition 1.3. *The vector fields $X_0, X_1, \dots, X_n$ are free up to weight s at 0 if and only if for any family of constants $\{c_I\}_{|I|\le s} \subset \mathbb{R}$ there exists a function $u \in C^\infty(\mathbb{R}^p)$ such that $X_I u(0) = c_I$ when $|I| \le s$.*

Proof. To show that the "if" condition holds, let us suppose that

$$\sum_{|I|\le s} a_I \left(X_{[I]}\right)_0 = 0.$$

Then

$$0 = \sum_{|I|\le s} a_I X_{[I]} u(0) = \sum_{|I|\le s} a_I \sum_{|J|\le s} A_{IJ} X_J u(0) = \sum_{|I|,|J|\le s} a_I A_{IJ} c_J,$$

and this implies that $\sum_{|I|\le s} a_I A_{IJ} = 0$ when $|J| \le s$, since the c_J's are arbitrary. Thus $X_0, X_1, \dots, X_n$ are free of weight s at 0.

Now we need some notation to prove the "only if" condition. We consider polynomials in the noncommuting variables $\xi_0, \xi_1, \dots, \xi_n$ and we assign to ξ_0 the weight $p_0 = 2$, and to ξ_i, for $i = 1, \dots, n$ the weight $p_i = 1$. As we did with the vector fields $X_0, X_1, \dots, X_n$, for any multi-index $I = (i_1, i_2, \dots, i_k)$ we put $\xi_{[I]} = \operatorname{ad}\xi_{i_1} \cdots \operatorname{ad}\xi_{i_{k-1}} \xi_{i_k}$, where $\operatorname{ad}\xi_i \xi_j = \xi_i\xi_j - \xi_j\xi_i$. Finally, we let V be the vector space spanned by the monomials ξ_I, with $|I| \le s$, and V' be its dual space. Every function $u \in C^\infty(\mathbb{R}^p)$ gives rise to the linear map $\Lambda_u \in V'$ defined by $\Lambda_u(p) = p(X_0, \dots, X_n)u(0)$, where $p \in V$. (Notation: if $p = \xi_I$, then $p(X_0, \dots, X_n)u(0) = (X_I u)(0)$). Thus we have a mapping

$$\Lambda : C^\infty(\mathbb{R}^p) \to V'$$
$$\Lambda : u \mapsto \Lambda_u$$

and our aim is to show that it is surjective. More precisely, if $L \in V'$ is defined by $L(\xi_I) = c_I$, we have to find $u \in C^\infty(\mathbb{R}^p)$ such that $\Lambda_u = L$. Let us denote with V_j the subspace of V spanned by the products $\xi_{[I_1]} \cdots \xi_{[I_\nu]}$, with $\nu \le j$ (and $|I_1| + \cdots + |I_\nu| \le s$). Notice that $V_s = V$. We will show by induction with respect to j, with $1 \le j \le s$, that there exists $u \in C^\infty(\mathbb{R}^p)$ such that $\Lambda_u = L$ on V_j, that is to say,

$$X_{[I_1]} \cdots X_{[I_\nu]} u(0) = L(\xi_{[I_1]} \cdots \xi_{[I_\nu]}) \tag{1.2}$$

if $\nu \le j$ and $|I_1| + \cdots + |I_\nu| \le s$.

If $j = 1$, then $\nu = 1$ and (1.2) can be written as

$$X_{[I]} u(0) = L(\xi_{[I]}) \quad \text{for any } |I| \le s.$$

Since the X_i's are free of weight s at 0, we have that

$$\sum_{|I| \le s} a_I (X_{[I]})_0 = 0 \;\Rightarrow\; \sum_{|I| \le s} a_I \xi_{[I]} = 0. \tag{1.3}$$

Namely, $\sum_{|I| \le s} a_I (X_{[I]})_0 = 0$ implies that $\sum_{|I| \le s} a_I A_{IJ} = 0$ for any J, hence

$$\sum_{|I| \le s} a_I \xi_{[I]} = \sum_{|I| \le s} a_I \sum_J A_{IJ} \xi_J = \sum_J \Big(\sum_{|I| \le s} a_I A_{IJ} \Big) \xi_J = 0.$$

By (1.3), there is a (unique) linear form defined on the span of the tangent vectors $\{(X_{[I]})_0\}_{|I| \le s}$ by

$$(X_{[I]})_0 \mapsto L(\xi_{[I]}).$$

We can extend this form to $\mathbb{R}^p$ and then find a function $u \in C^\infty(\mathbb{R}^p)$, e.g., a first degree homogeneous polynomial, such that the differential $u^{(1)}(0) = \mathrm{d}u(0)$ of u at 0 coincides with such an extension. Since $u^{(1)}(0)(X_{[I]})_0 = X_{[I]} u(0)$, the case $j = 1$ is done.

Assume now that for any $L \in V'$ there exists $u_0 \in C^\infty(\mathbb{R}^p)$ such that $\Lambda_{u_0} = L$ on V_{j-1}. This means that

$$X_{[I_1]} \cdots X_{[I_\nu]} u_0(0) = L(\xi_{[I_1]} \cdots \xi_{[I_\nu]}) \tag{1.4}$$

when $\nu \le j-1$ and $|I_1| + \cdots + |I_\nu| \le s$. If $u = u_0 + v$, with v vanishing of order j at 0 (in the usual sense), then $\Lambda_u = L$ on V_{j-1}, meaning that we must find v in such a way that (1.2) is solved for $\nu = j$. In this case, the equation takes the form

$$\text{(1.5)} \quad v^{(j)}(0)\Big((X_{[I_1]})_0, \ldots, (X_{[I_j]})_0\Big) = L(\xi_{[I_1]} \cdots \xi_{[I_j]}) - X_{[I_1]} \cdots X_{[I_j]} u_0(0),$$

where $v^{(j)}(0)$ is the j-th differential of v at 0, seen as a j-linear form on $\mathbb{R}^p$. Namely,

$$\begin{aligned} X_{[I_1]} \cdots X_{[I_j]} u(0) &= X_{[I_1]} \cdots X_{[I_j]} u_0(0) + X_{[I_1]} \cdots X_{[I_j]} v(0) \\ &= L(\xi_{[I_1]} \cdots \xi_{[I_\nu]}) \end{aligned}$$

but $X_{[I_1]} \cdots X_{[I_j]} v(0)$ simply equals $v^{(j)}(0)((X_{[I_1]})_0, \ldots, (X_{[I_j]})_0)$, because all the derivatives of v of intermediate order (which appear expanding the differential operator $X_{[I_1]} \cdots X_{[I_j]}$) actually vanish because v vanishes of order j at 0.

Thus, if we show that the right-hand side of (1.5) defines a symmetric j-linear form on the span of the tangent vectors $(X_{[I]})_0$, where $|I| \le s$, then we are done, because we can then extend this form to $\mathbb{R}^p$ and therefore find a function $v \in C^\infty(\mathbb{R}^p)$ vanishing of order j at 0, e.g., a j-th degree homogeneous polynomial, such that its j-th differential $v^{(j)}(0)$ at 0 coincides with the extended j-linear form.

So, let J be the form defined by

$$J : \Big((X_{[I_1]})_0, \ldots, (X_{[I_j]})_0\Big) \mapsto L(\xi_{[I_1]} \cdots \xi_{[I_j]}) - X_{[I_1]} \cdots X_{[I_j]} u_0(0).$$

The check that this j-linear form is actually well defined amounts to show that

$$\begin{aligned} &\sum a_{I_1}(X_{[I_1]})_0 = 0,\ \sum a_{I_2}(X_{[I_2]})_0 = 0,\ \ldots,\ \sum a_{I_j}(X_{[I_j]})_0 = 0 \\ &\qquad \Rightarrow \sum a_{I_1} a_{I_2} \ldots a_{I_j} \Big\{ L(\xi_{[I_1]} \cdots \xi_{[I_j]}) - X_{[I_1]} \cdots X_{[I_j]} u_0(0) \Big\} = 0. \end{aligned}$$

This is almost the same as in the $j = 1$ case. Indeed, the implication (1.3) still holds, and therefore $\sum a_{I_i}(X_{[I_i]})_0 = 0 \Rightarrow \sum a_{I_i} \xi_{[I_i]} = 0$ for $i = 1, 2, \ldots, n$; hence

$$\begin{aligned} &\sum a_{I_1} a_{I_2} \ldots a_{I_j} \Big\{ L(\xi_{[I_1]} \cdots \xi_{[I_j]}) - X_{[I_1]} \cdots X_{[I_j]} u_0(0) \Big\} \\ &= L\Big(\sum a_{I_1} \xi_{[I_1]} \sum a_{I_2} \xi_{[I_2]} \cdots \sum a_{Ij} \xi_{[Ij]}\Big) - \sum a_{I_1} X_{[I_1]} \sum a_{I_2} X_{[I_2]} \cdots \sum a_{Ij} X_{[I_j]} u_0(0) \\ &= 0. \end{aligned}$$

To show the symmetry of J, let us introduce

$$d_{I_1,\ldots,I_j} = L(\xi_{[I_1]} \cdots \xi_{[I_j]}) - X_{[I_1]} \cdots X_{[I_j]} u_0(0),$$

and let us prove that they are symmetric in the (multi-)indices. We first need to show that for every pair of multi-indices I and J, one has

$$[\xi_{[I]}, \xi_{[J]}] = \sum_{|K|=|I|+|J|} b_K \xi_{[K]},$$

where the b_K's are absolute constants, only depending on the multiindices I, J, K. This is just a consequence of Jacobi's identity, as we can show by induction on $\ell(I)$. First, if $\ell(I) = 1$, that is $I = (i)$, there is nothing to prove, because

$$[\xi_{[I]}, \xi_{[J]}] = [\xi_i, \xi_{[J]}] = \xi_{[K]}$$

with $K = (i, J)$, just by definition of $\xi_{[K]}$. Assume then the property for $\ell(I) \le k$, and let $I = (i, I')$ with $\ell(I') = k$; then

$$\begin{aligned} [\xi_{[I]}, \xi_{[J]}] &= [\xi_{[i,I']}, \xi_{[J]}] = [[\xi_i, \xi_{[I']}], \xi_{[J]}] \\ &= [\xi_i, [\xi_{[I']}, \xi_{[J]}]] + [\xi_{[I']}, [\xi_{[J]}, \xi_i]]. \end{aligned}$$

Now the first term in the last sum is already in the proper form, while the second can be rewritten in the proper form by inductive assumption, so we are done.

Let us show now the desired symmetry result. It is clearly sufficient to show it for consecutive indices. We limit ourselves to verify the symmetry with respect to the first two indices, the other cases being a straightforward generalization of it. Indeed, we have that

$$\begin{aligned} & d_{I_1,I_2,\dots,I_j} - d_{I_2,I_1,\dots,I_j} \\ &\quad = L([\xi_{[I_1]}, \xi_{[I_2]}] \cdots \xi_{[I_j]}) - [X_{[I_1]}, X_{[I_2]}] \cdots X_{[I_j]} u_0(0) \\ &\quad = \sum_{|K|=|I|+|J|} b_K \Big(L(\xi_{[K]} \xi_{[I_3]} \cdots \xi_{[I_j]}) - X_{[K]} X_{[I_3]} \cdots X_{[I_j]} u_0(0) \Big) = 0 \end{aligned}$$

by the induction hypothesis (1.4). Now we are (almost) done, because we have shown that the right-hand side of (1.5) defines a symmetric j-linear form on the span of $\{(X_{[I]})_0)\}_{|I| \le s}$. As we did for $j = 1$, we can extend this form to $\mathbb{R}^p$ and then find a function $v \in C^\infty(\mathbb{R}^p)$ vanishing of order j at 0, e.g., a j-th degree homogeneous polynomial, such that its j-th differential $v^{(j)}(0)$ at 0 coincides with the extended j-linear form. This completes the proof. □

Proposition 1.4. *Let $X_0, X_1, \dots, X_n$ be free of weight $s - 1$ but not of weight s at 0. Then one can find vector fields $\tilde{X}_j$ in $\mathbb{R}^{p+1}$ of the form*

$$\tilde{X}_j = X_j + u_j(x) \frac{\partial}{\partial t} \quad (j = 0, 1, \dots, n) \tag{1.6}$$

with $u_j \in C^\infty(\mathbb{R}^p)$, such that

(i) *The $\tilde{X}_j$'s remain free of weight $s-1$;*
(ii) *For every $r \geq s$,*

$$\dim \left\langle (\tilde{X}_{[I]})_0 \right\rangle_{|I| \leq r} = \dim \left\langle (X_{[I]})_0 \right\rangle_{|I| \leq r} + 1$$

where the symbol $\langle Y_\alpha \rangle_{\alpha \in B}$ denotes the vector space spanned by the vectors $\{Y_\alpha : \alpha \in B\}$.

Proof. Let us show that condition (i) in the above statement holds for any choice of the functions $u_j(x)$. To see this, we first claim that (1.6) implies

$$\tilde{X}_{[I]} = X_{[I]} + u_I(x)\frac{\partial}{\partial t} \tag{1.7}$$

for any multiindex I and some $u_I \in C^\infty(\mathbb{R}^p)$. Namely, we can proceed by induction on $\ell(I)$. For $\ell(I) = 1$, this is just (1.6); assume (1.7) holds for $\ell(I) = j-1$. For $\ell(I) = j$, let $I = (i, J)$ for some $i = 0, 1, \ldots, n$ and $\ell(J) = j-1$. Then, by inductive assumption,

$$\begin{aligned} \tilde{X}_{[I]} = \tilde{X}_{[i,J]} = [\tilde{X}_i, \tilde{X}_{[J]}] &= \left[X_i + u_i(x)\frac{\partial}{\partial t}, X_{[J]} + u_J(x)\frac{\partial}{\partial t} \right] \\ &= [X_i, X_{[J]}] + (X_i u_J - X_{[J]} u_i)\frac{\partial}{\partial t} = X_{[I]} + u_I(x)\frac{\partial}{\partial t}. \end{aligned}$$

Next, we show that (1.7) implies that the $\tilde{X}_i$'s are free of weight $s-1$. If

$$\sum_{|I| \leq s-1} a_I (\tilde{X}_{[I]})_0 = 0$$

for some coefficients a_I, then by (1.7) we have

$$\begin{aligned} 0 &= \sum_{|I| \leq s-1} a_I \left(X_{[I]} + u_I(x)\frac{\partial}{\partial t} \right)_0 \\ &= \sum_{|I| \leq s-1} a_I (X_{[I]})_0 + \Big(\sum_{|I| \leq s-1} a_I u_I(0) \Big) \frac{\partial}{\partial t}. \end{aligned}$$

Since $\partial/\partial t$ is independent from the vectors $(X_{[I]})_0$, this implies that

$$\sum_{|I| \leq s-1} a_I u_I(0) = 0 \quad \text{and} \quad \sum_{|I| \leq s-1} a_I (X_{[I]})_0 = 0.$$

But the X_i's are free of weight $s-1$ at 0, hence

$$\sum_{|I| \leq s-1} a_I A_{IJ} = 0 \quad \text{for any } J \text{ with } |J| \leq s-1.$$

Therefore also the $\tilde{X}_i$'s are free of weight $s-1$ at 0.

We now show that it is possible to choose smooth functions u_j such that condition (ii) in the statement of this proposition holds. To show this, we will prove that there exist functions u_j and constants $\{a_I\}_{|I|\le s}$ such that:

$$\sum_{|I|\le s} a_I (X_{[I]})_0 = 0, \tag{1.8}$$

$$\sum_{|I|\le s} a_I (\tilde{X}_{[I]})_0 \neq 0. \tag{1.9}$$

From (1.8)–(1.9), condition (ii) will follow; namely,

$$0 \neq \sum_{|I|\le s} a_I (\tilde{X}_{[I]})_0 = \sum_{|I|\le s} a_I \Big((X_{[I]})_0 + u_I(0) \frac{\partial}{\partial t} \Big) = \Big(\sum_{|I|\le s} a_I u_I(0) \Big) \frac{\partial}{\partial t} = b \frac{\partial}{\partial t}$$

with $b \neq 0$, hence

$$\frac{\partial}{\partial t} = \sum_{|I|\le s} \frac{a_I}{b} (\tilde{X}_{[I]})_0$$

and this shows that

$$\Big\langle (\tilde{X}_{[I]})_0 \Big\rangle_{|I|\le r} = \Big\langle (X_{[I]})_0 \Big\rangle \oplus \Big\langle \frac{\partial}{\partial t} \Big\rangle,$$

which implies condition (ii).

To prove (1.8)–(1.9), we use our assumption on the X_i: since they are *not* free of weight s, there exist coefficients $\{a_I\}_{|I|\le s}$ such that (1.8) holds but

$$\sum_{|I|\le s} a_I A_{IJ} \neq 0 \quad \text{for some } J \text{ with } |J| \le s. \tag{1.10}$$

It remains to prove that there exist functions u_j such that (1.9) holds for these u_j's and a_I's. To determine these u_j's, let us examine the action of the vector field

$$\sum_{|I|\le s} a_I \tilde{X}_{[I]} = \sum_{|I|\le s} a_I \sum_{|J|\le s} A_{IJ} \tilde{X}_J$$

on the function t. For any J with $|J| \le s$, let us write $J = (J'j)$ for some $j = 0, 1, \dots, n$. Then

$$\tilde{X}_J t = \tilde{X}_{J'} \tilde{X}_j t = \tilde{X}_{J'} \Big[\Big(X_j + u_j \frac{\partial}{\partial t} \Big) t \Big] = \tilde{X}_{J'} u_j = X_{J'} u_j$$

since u_j does not depend on t. We then have:

$$\Big(\sum_{|I|\le s} a_I \tilde{X}_{[I]}(t) \Big)(0) = \sum_{|I|\le s} a_I \sum_{|J|\le s} A_{IJ} \sum_{j=0,\dots,n;\, J=(J'j)} (X_{J'} u_j)(0).$$

Since $J = (J'j)$,

$$|J'| = \begin{cases} |J| - 1 & \text{for } j = 1, 2, \ldots, n, \\ |J| - 2 & \text{for } j = 0, \end{cases}$$

hence, in any case, $|J| \le s$ implies $|J'| \le s - 1$. Since the X_i's are free of weight $s-1$ at 0, by Proposition 1.3 for any choice of constants $\{c_{J'}\}_{|J'| \le s-1}$ there exists a function $u \in C^\infty(\mathbb{R}^p)$ such that $(X_{J'}u)(0) = c_{J'}$. On the other hand, by (1.10), there exists a set of constants $\{c_J\}_{|J| \le s}$ such that

$$\sum_{|I| \le s} \sum_{|J| \le s} a_I A_{IJ} c_J \ne 0.$$

Setting $c_{J'}^j = c_J$ if $J = (J'j)$ and applying $n+1$ times Proposition 1.3 to the $n+1$ sets of constants $\{c_{J'}^j\}_{|J'| \le s-1}$, $j = 0, 1, 2, \ldots, n$, we find $u_0, u_1, \ldots, u_n$ such that

$$\Big(\sum_{|I| \le s} a_I \tilde{X}_{[I]}(t) \Big)(0) = \sum_{|I| \le s} \sum_{|J| \le s} a_I A_{IJ} c_J \ne 0.$$

Hence (1.9) holds. This completes the proof of the proposition. $\square$

Theorem 1.5 (Lifting). *Let $X_0, X_1, \ldots, X_n$ be vector fields satisfying Hörmander's condition of step r at $x = 0$, that is the vectors*

$$\Big\{ (X_{[I]})_0 \Big\}_{|I| \le r}$$

span $\mathbb{R}^p$. (This clearly implies that such property holds in a suitable neighborhood of 0). Then there exist an integer m and vector fields $\tilde{X}_k$ in $\mathbb{R}^{p+m}$, of the form

$$\tilde{X}_k = X_k + \sum_{j=1}^{m} u_{kj}(x, t_1, t_2, \ldots, t_{j-1}) \frac{\partial}{\partial t_j}$$

$(k = 0, 1, \ldots, n)$, where the u_{kj}'s are polynomials, such that the $\tilde{X}_k$'s are free of weight r and $\{(\tilde{X}_{[I]})_0\}_{|I| \le r}$ span $\mathbb{R}^{p+m}$.

This theorem has an obvious reformulation in any point $x_0 \in \mathbb{R}^p$, with the lifted vector fields defined in a neighborhood of $(x_0, 0) \in \mathbb{R}^{p+m}$.

Proof. Let $\{(X_{[I]})_0\}_{I \in B}$ be a basis of $\mathbb{R}^p$, for some set B of p multiindices of weight $\le r$. We claim that

$$\operatorname{rank}[A_{IJ}]_{I \in B, |J| \le r} \ge p,$$

because the p vectors $(A_{IJ})_{|J|\le r}$, with $I \in B$, are independent. Namely, if for some constants $\{a_I\}_{I\in B}$

$$\sum_{I\in B} a_I A_{IJ} = 0 \quad \text{for any } J \text{ with } |J| \le r,$$

then

$$\sum_{I\in B} a_I (X_{[I]})_0 = \sum_{I\in B} a_I \sum_{|J|\le r} A_{IJ}(X_J)_0 = \sum_{|J|\le r} \Big(\sum_{I\in B} a_I A_{IJ} \Big)(X_J)_0 = 0.$$

But $\{(X_{[I]})_0\}_{I\in B}$ is a basis, hence $\sum_{I\in B} a_I (X_{[I]})_0 = 0$ implies $a_I = 0$ for any $I \in B$.

The relation just proved means that

$$p \le \operatorname{rank}[A_{IJ}]_{|I|,|J|\le r} \equiv c(r,n), \tag{1.11}$$

an absolute constant only depending on r, n.

Now, let $s \le r$ be such that $X_0, X_1, \dots, X_n$ are free of weight $s-1$ but not of weight s, at 0. (If the X_i's were already free of weight r, there would be nothing to prove. We also agree to say that the X_i's are free of weight 0 if they are not free of weight 1). We can then apply Proposition 1.4 and build vector fields

$$\tilde{X}_j = X_j + u_j(x)\frac{\partial}{\partial t} \quad (j = 0, 1, \dots, n)$$

in $\mathbb{R}^{p+1}$, free of weight $s-1$ and such that

$$\dim \Big\langle (\tilde{X}_{[I]})_0 \Big\rangle_{|I|\le r} = \dim \Big\langle (X_{[I]})_0 \Big\rangle_{|I|\le r} + 1 = p + 1$$

(because by assumption the $\{(X_{[I]})_0\}_{|I|\le r}$ span $\mathbb{R}^p$). Hence the $\{(\tilde{X}_{[I]})_0\}_{|I|\le r}$ still span the whole space $\mathbb{R}^{p+1}$. Now: either the vector fields $\{(\tilde{X}_{[I]})_0\}_{|I|\le r}$ are free of weight r, and we are done, or the assumptions of Proposition 1.4 are still satisfied, and we can iterate our argument; in this case, by (1.11), condition $p+1 \le c(r,n)$ must hold. After a suitable finite number m of iterations, condition $p + m \le c(r,n)$ cannot hold anymore, and this means that the vector fields $\tilde{X}_j$ must be free of weight r. The iterative argument also shows that the u_{kj}'s are polynomials only depending on the variables $x, t_1, t_2, \dots, t_{j-1}$. □

2. APPROXIMATION OF FREE SMOOTH VECTOR FIELDS

In this section we carry out the second part of Rothschild-Stein's procedure, that is, the approximation of free vector fields by left invariant vector fields on a homogeneous group. Here we concentrate on *smooth* vector fields, while the nonsmooth theory will be treated in Section 3. We actually prove a somewhat more general result than that by Rothschild-Stein, in the line of [11]. In Section 2.5, we will also prove, for free vector fields, a ball-box theorem and the resulting estimate on the volume of metric balls, in the spirit of Nagel-Stein-Wainger's results.

By the lifting theorem (Theorem 1.5), starting from any system of smooth Hörmander's vector fields in $\mathbb{R}^p$ we can define new vector fields $\tilde{X}_0, \dots, \tilde{X}_n$ in a neighborhood of $0 \in \mathbb{R}^{p+m}$, free up to weight r at 0 and such that $\{\tilde{X}_{[I]}(0)\}_{|I| \le s}$ spans $\mathbb{R}^{p+m}$. Just to simplify notation, throughout this section we will keep calling X_i and $\mathbb{R}^p$ these lifted free vector fields and their underlying space, respectively.

Therefore, let now $X_0, X_1, \dots, X_n$ be a system of *smooth* Hörmander's vector fields in $\Omega \subset \mathbb{R}^p$, *free up to weight* r and satisfying Hörmander's condition of step r in Ω. Since the X_i's are free, it is possible to choose a set B of p multiindices I with $|I| \le r$, such that $\{X_{[I]}\}_{I \in B}$ is a basis of $\mathbb{R}^p$ at any point $x \in \Omega$. We assume this set B fixed once and for all.

2.1. Canonical coordinates and weights of vector fields. Let us recall the standard definition of *exponential of a vector field*. We set:

$$\exp(tX)(\bar{x}) = \varphi(t)$$

where φ is the solution to the Cauchy problem

$$\begin{cases} \varphi'(\tau) = X_{\varphi(\tau)}, \\ \varphi(0) = \bar{x}. \end{cases} \tag{2.1}$$

The point $\exp(tX)(\bar{x})$ is uniquely defined for $t \in \mathbb{R}$ small enough, as soon as X has Lipschitz continuous coefficients, by the classical Cauchy's theorem about existence and uniqueness for solutions to Cauchy problems. For a fixed $\Omega' \Subset \Omega$, a t-neighborhood of zero where $\exp(tX)(\bar{x})$ is defined can be found uniformly for $\bar{x}$ ranging in Ω'.

Equivalently, we can write

$$\exp(tX)(\bar{x}) = \phi(1)$$

where ϕ is the solution to the Cauchy problem

$$\begin{cases} \phi'(\tau) = tX_{\phi(\tau)}, \\ \phi(0) = \bar{x}. \end{cases}$$

Now, for any $\bar{x} \in \Omega$, let us introduce the set of local ("canonical") coordinates

$$\mathbb{R}^p \ni u \longmapsto \exp\Big(\sum_{I \in B} u_I X_{[I]}\Big)(\bar{x}), \tag{2.2}$$

defined for u in a suitable neighborhood of 0. Note that the Jacobian of the map $u \longmapsto x$, at $u = 0$, equals the matrix of the vector fields $\{(X_{[I]})_{\bar{x}}\}_{I\in B}$, therefore is nonsingular. This allows to define canonical coordinates in a suitable neighborhood $U(\bar{x})$ of $\bar{x}$.

Since the basis $\{(X_{[I]})_{\bar{x}}\}_{I\in B}$ depends continuously on the point $\bar{x}$, the radius of this neighborhood can be taken uniformly bounded away from zero for $\bar{x}$ ranging in a compact set.

Henceforth in this section, all the computation will be made with respect to this system of coordinates defined in a neighborhood of the point $\bar{x}$ (which has canonical coordinates $u = 0$).

Our aim is to establish some basic properties enjoyed by the vector fields $X_{[I]}$, if they are expressed with respect to canonical coordinates, in particular Theorem 2.4, which will be a key tool for the following.

We start with the following result:

Lemma 2.1. *If we express the vector fields $X_{[I]}$ with respect to the above coordinates u, then we have that*

$$\sum_{I\in B} u_I \frac{\partial}{\partial u_I} = \sum_{I\in B} u_I X_{[I]}. \tag{2.3}$$

(*In the following, we will also write e_I for $\partial/\partial u_I$.*)

Proof. We start noting that, if

$$Y = \sum_{I\in B} y_I(u)\frac{\partial}{\partial u_I} \quad \text{and} \quad Z = \sum_{I\in B} z_I(u)\frac{\partial}{\partial u_I}$$

are two vector fields such that

$$Z(u_J) = Y(u_J) \quad \text{for any } J \in B,$$

(that is, the vector fields act in the same way on the functions $u \mapsto u_J$) then $y_I(u) = z_I(u)$ for any $I \in B$, hence $Y = Z$. Therefore, it will be enough to show that

$$\Big(\sum_{I\in B} u_I X_{[I]}\Big)(u_J) = \Big(\sum_{I\in B} u_I \frac{\partial}{\partial u_I}\Big)(u_J).$$

Now, for any vector field Y,

$$Yf(x) = \frac{\mathrm{d}}{\mathrm{d}t}\big(f(\exp(tY)(\bar{x}))\big)_{/t=t_0} \quad \text{where } x = \exp(t_0 Y)(\bar{x}).$$

Hence, if $Y = \sum_{I\in B} u_I X_{[I]}$, then

$$\Big(\sum_{I\in B} u_I X_{[I]}\Big)(u_J) = \frac{\mathrm{d}}{\mathrm{d}t}\Big(u_J\Big(\exp\Big(t\sum_{I\in B} u_I X_{[I]}\Big)(\bar{x})\Big)\Big)_{/t=t_0}$$

(just by definition of the coordinates u_I)

$$= \frac{\mathrm{d}}{\mathrm{d}t}(tu_J)_{/t=t_0} = u_J = \Big(\sum_{I\in B} u_I \frac{\partial}{\partial u_I}\Big)(u_J). \qquad \square$$

Definition 2.2 (Weights). We now assign the weight $|I|$ to the coordinate u_I and the weight $-|I|$ to $\partial/\partial u_I$. (Note that this is the convention made in [11], and is *different* from that made in [7] and [14]: in the last two papers, the authors assign positive weight also to derivatives). In the following we will say that a C^∞ function f has weight $\geqslant s$ if the Taylor expansion of f at the origin does not include terms of the kind $au_{I_1}u_{I_2}\cdots u_{I_k}$ with $a \neq 0$ and $|I_1|+|I_2|+\cdots+|I_k| < s$. A vector field $Y = \sum_{I\in B} f_I e_I$ has weight $\geqslant s$ if f_I has weight $\geqslant s+|I|$ for every $I \in B$.

Note that the weight of a function is always ≥ 0, while the weight of a vector field is $\geq -r$, where r is as above.

We want to stress that the definition of weight relies on the canonical coordinates, therefore it depends on the choice of a particular basis B of $\mathbb{R}^p$.

In the following we shall denote with F_s^q the set of functions such that in their Taylor expansion of degree $\leqslant q$ (in the standard sense), all terms have weight $\geq s$. Also V_s^q will denote the set of the vector fields with a similar property. The subset of F_s^q and V_s^q of elements that vanish at $u=0$ will be denoted by $\mathring{F}_s^q$ and $\mathring{V}_s^q$.

Lemma 2.3. *The following inclusions hold:*

$$
\begin{array}{lll}
F_s^q F_t^q \subset F_{s+t}^q, & \mathring{F}_s^q F_t^{q-1} \subset \mathring{F}_{s+t}^q, & \\
F_s^q V_t^q \subset V_{s+t}^q, & \mathring{F}_s^q V_t^{q-1} \subset \mathring{V}_{s+t}^q F_s^{q-1}, & \mathring{V}_t^q \subset \mathring{V}_{s+t}^q, \\
V_s^{q-1}(F_s^q) \subset V_{s+t}^{q-1}, & \mathring{V}_s^q(F_t^q) \subset \mathring{F}_{s+t}^q, & \\
[V_s^q, V_t^q] \subset V_{s+t}^{q-1}, & [\mathring{V}_s^q, V_t^{q-1}] \subset V_{s+t}^{q-1}, &
\end{array}
$$

with the obvious meaning of the symbols.

Proof. If $f \in F_s^q$ and $g \in F_t^q$, then all terms of the product of their Taylor expansion of degree $\leqslant q$ have weight $\geq s+t$. Therefore, the same is true for the Taylor expansion of degree $\leqslant q$ of fg, so that $fg \in F_{s+t}^q$. This shows that the first inclusion holds, and the second one is an immediate corollary. The other inclusions can be proved by means of similar arguments. $\square$

For any vector field $X_{[J]}$ with $|J| \leq s$, we can express $X_{[J]}$ in terms of the basis $\{X_{[I]}\}_{I\in B}$, writing

$$X_{[J]} = \sum_{I\in B} c_{JI}(u) X_{[I]}$$

for suitable functions c_{JI}.

Theorem 2.4. *For every multiindices I the vector field $X_{[I]}$ has weight $\geqslant -|I|$.*

Proof. Throughout the proof we will assume the $X_{[I]}$ written in canonical coordinates. In particular, the point $\bar{x}$ corresponds to $u = 0$. We shall prove by induction on $q \geqslant 0$ the following two facts:

(i) For every multiindex I we have $X_{[I]} \in V^q_{-|I|}$;
(ii) For every positive integer α, if for every $m \leqslant q+1$ and for every multiindices $I_1, \dots I_m \in B$ such that $|I_1| + \dots |I_m| < \alpha$ we have

$$X_{[I_m]} \cdots X_{[I_1]} f(0) = 0,$$

then $f \in F^{q+1}_\alpha$.

First of all we observe that it is enough to show that $X_{[I]} \in V^q_{-|I|}$ when $I \in B$. Indeed let us fix a certain $q \geqslant 0$ and assume to know that $X_{[I]} \in V^q_{-|I|}$ for every $I \in B$ and that (ii) holds. Let J be a multiindex, $J \notin B$. Then

$$X_{[J]} = \sum_{I \in B} c_{JI} X_{[I]}.$$

Since the vector fields $X_0, X_1, \dots, X_n$ are free up to step r, we can assume that in the last sum c_{JI} is nonzero only if $|I| = |J|$. For these constants c_{JI} we then have

$$X_{[J]} = \sum_{I \in B,\, |I|=|J|} c_{JI} X_{[I]},$$

which shows that it is enough to prove (i) for $I \in B$.

Let now $q = 0$. Observe that composing the operator $\operatorname{ad} e_I$ with (2.3) we get

$$X_{[I]} + \sum_{K \in B} u_K \operatorname{ad} e_I (X_{[K]}) = e_I. \tag{2.4}$$

This implies $(X_{[I]})_0 = e_I$ and therefore that $X_{[I]} \in V^0_{-|I|}$. Assume now that $f(0) = 0$ and that for every multiindex $I \in B$, $|I| < \alpha$, we have $X_{[I]} f(0) = 0$. Since $X_{[I]} f(0) = (\partial f / \partial u_I)(0)$ we have $f \in F^1_\alpha$.

Assume now that (i) and (ii) hold for a certain q and let us prove that the same is true with q replaced by $q+1$. We start with (i). We claim that it is enough to show that

$$W = \operatorname{ad} X_{[J]}(e_I) \in V^q_{-|I|-|J|}.$$

Indeed, $u_K \operatorname{ad} e_I(X_{[K]}) \in V^{q+1}_{-|I|}$ by Lemma 2.3, and since $e_I \in V^{q+1}_{-|I|}$ by (2.4) we have $X_{[I]} \in V^{q+1}_{-|I|}$.

In order to show that $W \in V^q_{-|I|-|J|}$ we compose $\operatorname{ad} X_{[J]}$ with (2.4). This yields

$$\text{(2.5)} \qquad \operatorname{ad} X_{[J]} e_I = W =$$

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

From (2.4) we also get $X_{[J]}(u_K) = \delta_{JK} - \sum_{L \in B} u_L \operatorname{ad} e_J(X_{[L]})(u_K)$. Since $\operatorname{ad} e_J(X_{[L]}) \in V^{q-1}_{-|J|-|L|}$ we have $u_L \operatorname{ad} e_J(X_{[L]}) \in \mathring{V}^q_{-|J|}$ and therefore

$$u_L \operatorname{ad} e_J(X_{[L]})(u_K) \in \mathring{F}^q_{-|J|+|K|}.$$

This implies that the second summation in (2.5) is congruent to $\operatorname{ad} e_I(X_{[J]}) = -W$ modulo $\mathring{V}^q_{-|J|-|I|}$. By Jacobi's identity (see the proof of Proposition 1.3)

$$\operatorname{ad} X_{[J]} X_{[I]} = \sum_{|L|=|J|+|I|} c_L X_{[L]}$$

for suitable coefficients c_L. This implies $\operatorname{ad} X_{[J]} X_{[I]} \in V^q_{-|J|-|I|}$. Hence

$$\text{(2.6)} \qquad W \equiv -W + \sum_{K \in B} u_K \operatorname{ad} X_{[J]} \operatorname{ad} e_I(X_{[K]}) \pmod{V^q_{-|J|-|I|}}.$$

Since

$$\begin{aligned} u_K \operatorname{ad} X_{[J]} \operatorname{ad} e_I(X_{[K]}) &= -u_K \operatorname{ad} X_{[J]} \operatorname{ad} X_{[K]}(e_I) \\ &= -u_K \operatorname{ad} X_{[K]} \operatorname{ad} X_{[J]}(e_I) - u_K \operatorname{ad}[X_{[K]}, X_{[J]}](e_I) \end{aligned}$$

we have

$$u_K \operatorname{ad} X_{[J]} \operatorname{ad} e_I(X_{[K]}) \equiv -u_K \operatorname{ad} X_{[K]} \operatorname{ad} X_{[J]}(e_I) \pmod{V^q_{-|I|-|J|}}.$$

By Jacobi's identity and substituting in (2.6) we obtain

$$2W = -\sum_K u_K \operatorname{ad} X_{[K]}(W) \pmod{V^q_{-|I|-|J|}}.$$

We now use (2.3) to replace $X_{[K]}$ by e_K. Indeed we have

$$\begin{aligned} \sum_K u_K \operatorname{ad} X_{[K]}(W) &= \sum_K \operatorname{ad}(u_K X_{[K]})(W) + \sum_K W(u_k) X_{[K]} \\ &= \operatorname{ad}\Big(\sum_K u_K X_{[K]}\Big)(W) + \sum_K W(u_k) X_{[K]} \\ &= \operatorname{ad}\Big(\sum_K u_K e_K\Big)(W) + \sum_K W(u_k) X_{[K]} \\ &= \sum_K u_K \operatorname{ad}(e_K)(W) + \sum_K W(u_K)(X_{[K]} - e_K). \end{aligned}$$

Since $W = \operatorname{ad} X_{[J]} e_I \in V^{q-1}_{-|I|-|J|}$, we have $W(u_K) \in F^{q-1}_{|K|-|I|-|J|}$. Also, since $X_{[K]} - e_K \in \mathring{V}^q_{-|K|}$ we have $W(u_K)(X_{[K]} - e_K) \in V^q_{-|I|-|J|}$. Hence

$$TW \in V^q_{-|I|-|J|}$$

where we set $TW = 2W + \sum_K u_K \operatorname{ad}(e_K)(W)$. We claim that this implies that $W \in V^q_{-|I|-|J|}$. To see this, write $W = \sum_L f_L e_L$. Then

$$\begin{aligned} TW &= 2\sum_L f_L e_L + \sum_{L,K} u_K [e_K, f_L e_L] \\ &= 2\sum_L f_L e_L + \sum_{L,K} u_K \frac{\partial f_L}{\partial u_K} e_L \\ &= \sum_L \left(2f_L + \sum_K u_K \frac{\partial f}{\partial u_K}\right) e_L. \end{aligned}$$

Let g be a homogeneous function of degree μ, then $\sum_K u_K \, \partial g/\partial u_K = \mu g$ which shows that the operator $f \mapsto 2f + \sum_K u_K \, \partial f/\partial u_K$ acts on the Taylor expansion of a function multiplying a term of degree μ by $(2+\mu)$. This implies that $W \in V^q_{-|I|-|J|}$.

Now we will show that also (ii) holds with q replaced by $q+1$. We have to show that if for every $m \leqslant q+2$ and for every multiindices $I_1, \dots I_m \in B$ such that $|I_1| + \dots |I_m| < \alpha$ we have $X_{[I_m]} \cdots X_{[I_1]} f(0) = 0$, then $f \in F^{q+2}_\alpha$. By the definition of the class F^{q+2}_α this amounts to showing that $e_{I_m} \cdots e_{I_1} f(0) = 0$ for every $m \leqslant q+2$ and every $I_1, \dots I_m \in B$ such that $|I_1| + \dots |I_m| < \alpha$.

By the induction hypothesis $f \in F^{q+1}_\alpha$, we already know that for every $I \in B$ we have $X_{[I]} \in V^{q+1}_{-|I|}$. By (2.4) we have

$$e_{I_1} f = X_{[I_1]} f + g_1$$

with $g_1 = \sum_{K \in B} u_K \operatorname{ad} e_{I_1}(X_{[K]}) f$. By Lemma 2.3 we have $g_1 \in F^{q+1}_{\alpha-|I_1|}$. Iterating this argument yields

$$e_{I_m} \dots e_{I_1} f = X_{[I_m]} \cdots X_{[I_1]} f + g_m$$

with $g_m \in F^{q+2-m}_{\alpha-(|I_1|+\cdots+|I_m|)}$. Since $\alpha > |I_1| + \cdots + |I_m|$ this implies $g_m(0) = 0$ and therefore that $e_{I_m} \dots e_{I_1} f(0) = 0$. ◻

2.2. Pointwise approximation. As in the previous subsection, we assume that the X_i's are free and a basis $\{(X_{[I]})_{\bar{x}}\}_{I \in B}$ for $\mathbb{R}^p$ is chosen once and for all; this choice induces a system of canonical coordinates u near $\bar{x}$ such that

$$\sum_{I \in B} u_I e_I = \sum_{I \in B} u_I X_{[I]} \tag{2.7}$$

(where $e_I = \partial/\partial u_I$) and that for every multiindex I the vector field $X_{[I]}$ has weight $\geq -|I|$.

We can now prove Rothschild-Stein's approximation theorem for free weighted vector fields:

Theorem 2.5 (Approximation, pointwise version). *If $Y_0, Y_1, \dots, Y_n$ is another system of vector fields satisfying (with respect to the same canonical coordinates)*

$$\sum_{I\in B} u_I e_I = \sum_{I\in B} u_I Y_{[I]}, \tag{2.8}$$

then $X_{[I]} - Y_{[I]}$ has weight $\geq 1 - |I|$.

(In particular, for $I = (i)$ we have that $X_0 - Y_0$ has weight ≥ -1 while $X_i - Y_i$ has weight ≥ 0 for $i = 1, 2, \dots, n$).

Proof. The proof exploits the same techniques as in the proof of Theorem 2.4. Let us recall that we are now working in canonical coordinates. We shall prove by induction on q that

$$X_{[I]} - Y_{[I]} \in V^q_{1-|I|}. \tag{2.9}$$

Observe that this is obvious when $|I| > s$. Indeed, every vector field $Z = \sum_{J\in B} f_J e_J$ has a weight $\geqslant -s$ since if $J \in B$, then $|J| \leqslant s$. Next we prove that it is enough to show (2.9) when $I \in B$.

Indeed, let I be any multiindex of weight $\leqslant s$. Then since $\{(X_{[J]})_{\bar{x}}\}_{J\in B}$ spans $\mathbb{R}^p$, there exist coefficients c_{IJ} such that

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

Let $a_J = \delta_{IJ} - c_{IJ}$ where we assume that $c_{IJ} = 0$ if $J \notin B$. Then

$$\sum a_J (X_{[J]})_{\bar{x}} = 0.$$

Since the vector fields are free of weight s, this implies that for every K

$$\sum_J a_J A_{JK} = 0.$$

Therefore, for any family of vector fields Z_i we have

$$0 = \sum_{K, |K|=|I|} \sum_J a_J A_{JK} Z_K = \sum_{J, |J|=|I|} \sum_K a_J A_{JK} Z_K = \sum_{|J|=|I|} a_J Z_{[J]}$$

and finally

$$Z_{[I]} = \sum_{J\in B, |J|=|I|} c_{IJ} Z_{[J]}.$$

In particular we have

$$X_{[I]} = \sum_{J\in B,\, |J|=|I|} c_{IJ} X_{[J]}$$

and

$$Y_{[I]} = \sum_{J\in B,\, |J|=|I|} c_{IJ} Y_{[J]}.$$

These identities show that if (2.9) hold for $I \in B$, then they hold for every I.

Now we prove that (2.9) hold for $I \in B$. Let $q = 0$. We have seen in the proof of Theorem 2.4 that (2.8) implies

$$X_{[I]} + \sum_{K\in B} u_K \operatorname{ad} e_I(X_{[K]}) = e_I \tag{2.10}$$

and

$$Y_{[I]} + \sum_{K\in B} u_K \operatorname{ad} e_I(Y_{[K]}) = e_I. \tag{2.11}$$

In particular, $(X_{[I]})_{\bar{x}} = e_I$ and $(Y_{[I]})_{\bar{x}} = e_I$, hence (2.9) holds for $q = 0$.

We now assume that (2.9) holds for a certain q and we prove that the same is true with q replaced by $q + 1$. We claim that it is enough to show that

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

Indeed, by (2.10), (2.11)

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

and by Lemma 2.3, $u_K \operatorname{ad}(X_{[K]} - Y_{[K]})(e_I) \in V^{q+1}_{1-|I|}$.

Now we will show that $Z = \operatorname{ad}(X_{[J]} - Y_{[J]})(e_I) \in V^q_{1-|I|-|J|}$. Composing (2.10) with $\operatorname{ad} X_{[J]}$, (2.11) with $\operatorname{ad} Y_{[J]}$ and computing the difference gives

$$\begin{aligned}
Z &= \operatorname{ad} X_{[J]}(X_{[I]}) - \operatorname{ad} Y_{[J]}(Y_{[I]}) + \sum_{K\in B} \operatorname{ad} X_{[J]}(u_K \operatorname{ad} e_I(X_{[K]})) \\
&\qquad - \sum_{K\in B} \operatorname{ad} Y_{[J]}(u_K \operatorname{ad} e_I(Y_{[K]})) \\
&= \operatorname{ad} X_{[J]}(X_{[I]}) - \operatorname{ad} Y_{[J]}(Y_{[I]}) + \sum_{K\in B} \operatorname{ad}(X_{[J]} - Y_{[J]})(u_K \operatorname{ad} e_I(X_{[K]})) \\
&\qquad + \sum_{K\in B} \operatorname{ad} Y_{[J]}(u_K \operatorname{ad} e_I(X_{[K]} - Y_{[K]})).
\end{aligned}$$

By the Jacobi identity

$$\operatorname{ad} X_{[J]}(X_{[I]}) - \operatorname{ad} Y_{[J]}(Y_{[I]}) = \sum_{|K|=|J|+|I|} c_K (X_{[K]} - Y_{[K]})$$

and by inductive hypothesis $\operatorname{ad} X_{[J]}(X_{[I]}) - \operatorname{ad} Y_{[J]}(Y_{[I]}) \in V^q_{1-|I|-|J|}$.

Also, since $X_{[K]} \in V^{q+1}_{-|K|}$, we have $\operatorname{ad} e_I(X_{[K]}) \in V^q_{-|I|-|K|}$, and therefore $u_K \operatorname{ad} e_I(X_{[K]}) \in \mathring{V}^{q+1}_{-|I|}$. Since $X_{[J]} - Y_{[J]} \in V^q_{1-|J|}$ we have

$$\operatorname{ad}(X_{[J]} - Y_{[J]})(u_K \operatorname{ad} e_I(X_{[K]})) \in V^q_{1-|I|-|J|}.$$

This means that modulo $V^q_{1-|I|-|J|}$ we have

$$\begin{aligned} Z &\equiv \sum_{K\in B} \operatorname{ad} Y_{[J]}(u_K \operatorname{ad} e_I(X_{[K]} - Y_{[K]})) \\ &\equiv \sum_{K\in B} \operatorname{ad}(Y_{[J]} - e_J)(u_K \operatorname{ad} e_I(X_{[K]} - Y_{[K]})) + \sum_{K\in B} \operatorname{ad} e_J(u_K \operatorname{ad} e_I(X_{[K]} - Y_{[K]})) \\ &\equiv \sum_{K\in B} \operatorname{ad} e_J(u_K \operatorname{ad} e_I(X_{[K]} - Y_{[K]})) \end{aligned}$$

since $Y_{[J]} - e_J \in \mathring{V}^{q+1}_{-|J|}$ and $u_K \operatorname{ad} e_I(X_{[K]} - Y_{[K]}) \in V^q_{1-|I|}$. Finally

$$\begin{aligned} Z &\equiv \sum_{K\in B} \operatorname{ad} e_J(u_K \operatorname{ad} e_I(X_{[K]} - Y_{[K]})) \\ &\equiv \sum_{K\in B} [\operatorname{ad} e_J \operatorname{ad} e_I[u_K(X_{[K]} - Y_{[K]})] - \operatorname{ad} e_J \delta_{IK}(X_{[K]} - Y_{[K]})] \\ &\equiv \operatorname{ad} e_J \operatorname{ad} e_I\Big[\sum_{K\in B} u_K(X_{[K]} - Y_{[K]})\Big] - \operatorname{ad} e_J(X_{[I]} - Y_{[I]}) \\ &\equiv -\operatorname{ad} e_J(X_{[I]} - Y_{[I]}) \equiv \operatorname{ad}(X_{[I]} - Y_{[I]})(e_J), \end{aligned} \tag{2.12}$$

since $\sum_{K\in B} u_K(X_{[K]} - Y_{[K]}) = 0$ by (2.7) and (2.8). This shows that for every multiindex I and J we have $\operatorname{ad}(X_{[J]} - Y_{[J]})(e_I) \equiv \operatorname{ad}(X_{[I]} - Y_{[I]})(e_J)$. Using this fact in (2.12) yields

$$\begin{aligned} Z &\equiv \sum_{K\in B} \operatorname{ad} e_J(u_K \operatorname{ad} e_I(X_{[K]} - Y_{[K]})) \\ &\equiv \sum_{K\in B} \operatorname{ad} e_J(u_K \operatorname{ad} e_K(X_{[I]} - Y_{[I]})) \\ &\equiv \operatorname{ad} e_J(X_{[I]} - Y_{[I]}) + \sum_{K\in B} u_K \operatorname{ad} e_J \operatorname{ad} e_K(X_{[I]} - Y_{[I]}) \\ &\equiv \operatorname{ad} e_J(X_{[I]} - Y_{[I]}) + \sum_{K\in B} u_K \operatorname{ad} e_K \operatorname{ad} e_J(X_{[I]} - Y_{[I]}) \end{aligned}$$

since e_K and e_J commute. Hence

$$Z \equiv -Z - \sum_{K\in B} u_K \operatorname{ad} e_K Z.$$

This means

$$TZ \equiv 0 \quad (\operatorname{mod} V^q_{1-|I|-|J|}).$$

which implies $Z \in V^q_{1-|I|-|J|}$. □

In order to recover from Theorem 2.5 the exact statement of Rothschild-Stein's "approximation theorem", some work has still to be done. First, we have to pass from the *pointwise* statement of Theorem 2.5 to an analogous *local* statement. This involves the introduction of Rothschild-Stein's "map Θ" and the study of some of its properties. Second, we have to apply this theorem to the case where the vector fields Y_i are homogeneous left invariant with respect to a structure of homogeneous group, and deduce some information on the "remainders" in this approximation procedure. These tasks will be performed in the next two subsections, respectively.

***2.3. From* pointwise *to* local. *The map* Θ.** We now revise the construction of local coordinates u_I made in Section 2.1. Let Ω be as at the beginning of Section 2; we claim that for any $\Omega' \Subset \Omega$ there exists a neighborhood $U(0) \subset \mathbb{R}^p$ where the map

$$E(\cdot, \xi_0) : u \equiv (u_I)_{I \in B} \longmapsto \xi \equiv \exp\Big(\sum_{I \in B} u_I X_{[I]}\Big)(\xi_0) \tag{2.13}$$

is well defined and smooth, for any fixed $\xi_0 \in \Omega'$. Namely, by classical results about O.D.E.'s, E is smooth in the joint variables $(u, \xi_0) \in U(0) \times \Omega'$.

Next, we define

$$F(u, \xi_0, \xi) = E(u, \xi_0) - \xi$$

on $U(0) \times \Omega' \times \mathbb{R}^p$. Noting that $F(0, \xi_0, \xi_0) = 0$ and that the Jacobian of F with respect to the u variables, at $(0, \xi_0, \xi_0)$, has determinant

$$\det\Big((X_{[I]})_{\xi_0}\Big),$$

which does not vanish since $\{X_{[I]}\}_{I \in B}$ span $\mathbb{R}^p$, by the implicit function theorem we can define a function

$$u = \Theta(\eta, \xi),$$

smooth in some neighborhood W of (ξ_0, ξ_0), such that $E(\Theta(\eta, \xi), \eta) = \xi$.

Summarizing the discussion above, we can state the following result:

Proposition 2.6 (The map Θ***).***

(i) *For any* $\xi_0 \in \Omega$ *there exist a neighborhood* W *of* (ξ_0, ξ_0) *in* $\mathbb{R}^{2p}$, *a neighborhood* $U(0)$ *of* 0 *in* $\mathbb{R}^p$ *and a smooth map* $\Theta(\cdot, \cdot) : W \to U(0)$ *such that:*

$$\xi = \exp\Big(\sum_{I \in B} u_I X_{[I]}\Big)(\eta) \quad \text{for } u = \Theta(\eta, \xi); \tag{2.14}$$

(ii) *The map* Θ *satisfies*

$$\Theta(\eta,\xi) = -\Theta(\xi,\eta); \tag{2.15}$$

(iii) *For any fixed* η*, the map* $u = \Theta(\eta,\xi)$ *is a diffeomorphism from a neighborhood of* η *onto a neighborhood of* 0*, in* $\mathbb{R}^p$*;*

(iv) *Analogously, for any fixed* ξ*, the map* $u = \Theta(\eta,\xi)$ *is a diffeomorphism from a neighborhood of* ξ *onto a neighborhood of* 0*.*

Proof. We have already proved (i) and (iii); (ii) follows form the fact that, for any vector field X,

$$\xi = \exp(X)(\eta) \;\Rightarrow\; \eta = \exp(-X)(\xi),$$

as can be checked by definition of the exponential map; (iv) is then a consequence of (ii) and (iii). □

The map Θ allows one to restate Theorem 2.5 (approximation) in a form more similar to that of Rothschild-Stein.

Recall that a vector field Z has weight k at some fixed point η if Z, expressed in terms of the local coordinates $u = \Theta(\eta,\xi)$, has weight k at $u = 0$, in the sense of Definition 2.2.

It will be useful to recall also the concrete meaning of expressing the same vector field in different coordinates: if we denote by Z^{ξ} and Z^{u}, the vector field Z written as a differential operator which acts on the variables ξ or u, respectively, then

$$Z^{\xi}[f(\Theta(\eta,\xi))] = (Z^{u}f)(\Theta(\eta,\xi)), \tag{2.16}$$

for any smooth function $f(u)$.

Then we have the following result:

Theorem 2.7 (Approximation, local version). *For every multiindex* I *the vector field* $X_{[I]}$ *has weight* $\geq -|I|$ *at any point of* Ω*. If* $Y_0, Y_1, \ldots, Y_n$ *is another system of vector fields (expressed in the same coordinates* u*) satisfying*

$$\sum_{I\in B} u_I e_I = \sum_{I\in B} u_I Y_{[I]}, \tag{2.17}$$

then $X_{[I]} - Y_{[I]}$ *has weight* $\geq 1-|I|$ *at any point of* Ω*. Moreover, for any point* $\eta\in\Omega$ *there exists a system of vector fields* $R_{\eta,[I]}$*, of weight* $\geq 1-|I|$ *at* η *(when expressed in the coordinates* u*) and smoothly depending on the point* η*, such that*

$$X^{\xi}_{[I]}[f(\Theta(\eta,\xi))] = (Y_{[I]}f)(\Theta(\eta,\xi)) + (R_{\eta,[I]}f)(\Theta(\eta,\xi)). \tag{2.18}$$

Proof. The first part of the theorem is exactly Theorem 2.4 plus Theorem 2.5, stated at any point η. Saying that $X_{[I]} - Y_{[I]}$ has weight $\geq 1 - |I|$ at η, just by definition means that the vector field $R_{\eta,[I]} = X^u_{[I]} - Y_{[I]}$ has weight $\geq 1 - |I|$ at $u = 0$. Here the superscript u in $X^u_{[I]}$ emphasizes that this vector field is expressed in terms of the coordinates u. By (2.16), we can rewrite it in terms of coordinates ξ, getting (2.18). It remains to check that $R_{\eta,[I]}$ depends smoothly on η. Let

$$R_{\eta,[I]} = \sum_J b_{IJ}(\eta, u)\,\partial_{u_J};$$

then, applying (2.18) to the function $f(u) = u_J$ we get

$$b_{IJ}(\eta, \Theta(\eta,\xi)) = X^{\xi}_{[I]}[(\Theta(\eta,\xi))_J] - (Y_{[I]}u_J)(\Theta(\eta,\xi)).$$

The right-hand side of this equation is a smooth function of (η,ξ), since Θ is smooth (see Proposition 2.6); hence the functions $(\eta,\xi) \longmapsto b_{IJ}(\eta,\Theta(\eta,\xi))$ are smooth; fixing ξ and composing with the diffeomorphism $u = \Theta(\eta,\xi)$ we read that $b_{IJ}(\eta,u)$ are smooth functions, which is what we needed to prove. □

Remark 2.8. The last statement is perfectly analogous to the approximation theorem proved by Rothschild-Stein, but somewhat more general, since the vector fields $Y_{[I]}$ need not be left invariant on a homogeneous group; they only need to satisfy (2.17).

2.4. Approximation by left invariant vector fields. The standard application of Theorem 2.7 requires the construction of a particular system of vector fields $\{Y_{[I]}\}_{I\in B}$ enjoying special properties.

In the following statement, the numbers p, n, r keep the same meaning they had in the previous subsections; also the system of multiindices $(I \in B)$ is the same.

Theorem 2.9. *There exist in $\mathbb{R}^p$ a system of smooth vector fields Y_0, Y_1, $\ldots$, Y_n and a structure of homogeneous group G, that is, a Lie group operation $u \circ v$ ("translation") and a one-parameter family $\{\delta_\lambda\}_{\lambda>0}$ of automorphisms ("dilations"), acting as*

$$\delta_\lambda((u_I)_{I\in B}) = (\lambda^{|I|}u_I)_{I\in B},$$

such that:

(i) *The vector fields $Y_0, Y_1, \ldots, Y_n$ are free up to weight r in $\mathbb{R}^p$ and the vectors $\{(Y_{[I]})_u\}_{|I|\leq r}$ span $\mathbb{R}^p$ at any point u of the space;*
(ii) *The $Y_{[I]}$'s are left invariant and homogeneous of degree $|I|$ with respect to the dilations in G;*
(iii) *At $u = 0$, the $Y_{[I]}$'s coincide with the local basis associated to the coordinates u_I, that is,*

$$(Y_{[I]})_0 = \frac{\partial}{\partial u_I};$$

(iv) *The $Y_{[I]}$'s satisfy* (2.17);

(v) *In the group G, the inverse u^{-1} of an element is just its (Euclidean) opposite $-u$.*

We stress that all the previous properties hold simultaneously, with respect to the same system of coordinates in the space $\mathbb{R}^p$. These coordinates will be identified with the canonical coordinates u induced by the vector fields $X_{[I]}$ (see Section 2.1).

Proof. For the following abstract construction we refer to [13, pp. 3-15].

1. Let us consider the Lie algebra $\mathfrak{g}$ obtained by quotienting the free Lie algebra with generators $Z_0, \dots, Z_n$ with respect to the ideal spanned by all commutators of weight greater than r (this is called the *free nilpotent Lie algebra of type II* in [14]); here $Z_0, \dots, Z_n$ are thought as abstract generators, having weight 2 (Z_0) and 1 ($Z_1, Z_2, \dots, Z_n$). This abstract Lie algebra is isomorphic to $\mathbb{R}^d$ for some d. We claim that actually $d = p$. Namely the structure of the free Lie algebra of type II of step r on n generators can depend only on n, r, and since the Lie algebra generated by the X_i's is $\mathbb{R}^p$, $p = d$.

Hence the Lie algebra $\mathfrak{g}$ will be identified with $\mathbb{R}^p$ from now on.

2. We then introduce in $\mathbb{R}^p$ an operation $\circ$, defined by:

$$x \circ y \equiv S(x,y) \equiv x + y + \frac{1}{2}[x,y] + \frac{1}{12}[[x,y],y] - \frac{1}{12}[[x,y],x] + \cdots \tag{2.19}$$

In the previous formula, $[x,y]$ denotes the commutator in the Lie algebra $\mathfrak{g}$ (whose elements have been identified with points of $\mathbb{R}^p$); the sum is finite because the Lie algebra is nilpotent, and the precise definition of S is given by

$$S(\cdot,\cdot) = S'(1,1,\cdot,\cdot)$$

where S' is the function appearing in the Baker-Campbell-Hausdorff formula:

$$\exp(sX)\exp(tY) = \exp(S'(s,t,X,Y)) \tag{2.20}$$

which holds, for s, t small enough, for any couple of smooth vector fields X, Y which generate a finite dimensional Lie algebra. More precisely, it is known that

$$S(x,y) = \sum_{j+k\geq 1} Z_{j,k}(x,y) \tag{2.21}$$

where each $Z_{j,k}(x,y)$ is a fixed linear combination of iterated commutators of x and y, containing j times x and k times y. In terms of coordinates in $\mathbb{R}^p$, this function can be written as

$$S(x,y) = (S_1(x,y), S_2(x,y), \dots, S_p(x,y))$$

where each S_j is a polynomial in x, y. Then (see Theorem 4.2 in [13]) the operation $\circ$ defines in $\mathbb{R}^p$ a structure of homogeneous Lie group G, whose Lie algebra $\mathrm{Lie}(G)$ is isomorphic to $\mathfrak{g}$.

3. The isomorphism of $\mathfrak{g}$ with $\mathrm{Lie}(G)$, explicitly, means that if we define $Y_{[I]}$ as the left invariant (with respect to G) vector field in $\mathbb{R}^p$ which agrees with ∂_{u_I} at the origin, then the Lie algebra generated by $\{Y_{[I]}\}_{I\in B}$ is isomorphic to $\mathfrak{g}$; in particular, it is free up to weight r and the vectors $\{(Y_{[I]})_u\}_{|I|\le r}$ span $\mathbb{R}^p$ at any point. Clearly, this isomorphism logically depends on the definition of $\circ$ in terms of the Baker-Campbell-Hausdorff formula.

4. It remains to show (iv) and (v). Both follow from taking a look inside the operation $S(x,y)$. As to (v), from (2.19) we read that

$$S(x,-x) = x - x + \frac{1}{2}[x,-x] + \frac{1}{12}[[x,-x],-x] - \frac{1}{12}[[x,-x],x] + \cdots = 0$$

since $[x,x] = 0$. Therefore the Euclidean opposite is also the inverse in the group. To prove (iv), we start writing, for any smooth function f,

$$\begin{aligned}
(Y_{[I]}f)(x) &= \frac{\mathrm{d}}{\mathrm{d}t}_{/t=0} f(x \circ te_I) \qquad \text{by (2.21)},\\
&= \frac{\mathrm{d}}{\mathrm{d}t}_{/t=0} f\Big(\sum_{j+k\ge 1} Z_{j,k}(x, te_I)\Big)\\
&= \nabla f(x) \cdot \sum_{j+k\ge 1} \frac{\mathrm{d}}{\mathrm{d}t}_{/t=0} t^k Z_{j,k}(x, e_I)\\
&= \nabla f(x) \cdot \sum_{j\ge 0} Z_{j,1}(x, e_I).
\end{aligned}$$

Then we compute

$$\begin{aligned}
\sum_{I\in B} x_I (Y_{[I]}f)(x) &= \sum_{I\in B} \nabla f(x) \cdot \sum_{j\ge 0} Z_{j,1}(x, x_I e_I)\\
&= \nabla f(x) \cdot \sum_{j\ge 0} Z_{j,1}(x,x)\\
&= \nabla f(x) \cdot x = \sum_{I\in B} x_I \, \partial_{x_I} f(x),
\end{aligned}$$

that is, (2.17). The theorem is completely proved. □

Theorem 2.7 can now be applied choosing the left invariant vector fields $Y_{[I]}$ as the approximating system. The map $u = \Theta(\eta,\xi)$ can now be regarded as a diffeomorphism from a neighborhood of η onto a neighborhood of 0 *in the group* G. In other words, $\Theta(\eta,\xi)$ is an element of the group G, and one has:

$$\Theta(\eta,\xi) = -\Theta(\xi,\eta) = \Theta(\xi,\eta)^{-1}.$$

Remark 2.10. Before stating their Theorem 5, Hörmander and Melin suggest how to connect it with Theorem 5 in [14]. We would like now to explain in more details this link, giving in this way a reformulation of some of the results of Theorem 2.9.

Let us consider the Lie algebra $\mathfrak{g}$ obtained by quotienting the free Lie algebra with generators $X_0, \ldots, X_n$ with respect to the ideal spanned by all commutators of weight greater than r. If G is the (unique, nilpotent) connected and simply connected Lie group having $\mathfrak{g}$ as its Lie algebra, then the exponential map exp : $\mathfrak{g} \to G$ is a diffeomorphism, so that $\mathbb{R}^p \simeq \mathfrak{g} \simeq G$. We denote with $Y_i \in \mathfrak{g}$ the equivalence class of X_i. It can be seen, as usual, as a left-invariant vector field on G. It is also clear that $\{(Y_{[I]})_0\}_{I\in B}$ is a basis of the tangent space T_0G, where 0 here stands for the identity of G. We can thus define a system of coordinates $(u_I)_{I\in B}$ on G by means of

$$\mathbb{R}^p \ni (u_I)_{I\in B} \mapsto \exp\Big(\sum_{I\in B} u_I Y_{[I]}\Big)(0) = \exp\Big(\sum_{I\in B} u_I Y_{[I]}\Big) \in G, \tag{2.22}$$

as we did previously for the vector fields X_i on $\mathbb{R}^p$. (The first exp in (2.22) is the exponential of a vector field, while the second one is the exponential map of the group G). Notice however that in this case the coordinate system is global, since the exponential map is a diffeomorphism. We used the same notation for the two sets of coordinates because they will be immediately identified. Indeed, we can associate with $(u_I)_{I\in B} \in \mathbb{R}^p$ a point in $\mathbb{R}^p$ by (2.2) and an element in G by (2.22). This provides a map from G to $\mathbb{R}^p$, and allows us to compare the vector fields $X_{[I]}$ and $Y_{[I]}$, once they have been written in these coordinates. Put in a little bit different way, we can use (2.22) to identify G with the "same" $\mathbb{R}^p$ where the vector fields X_i are defined, so that the X_i and the Y_i live in the same space. At this point, in order to apply Theorem 2.5 we simply have to notice that, arguing as in the proof of Lemma 2.3, one has

$$\sum_{I\in B} u_I e_I = \sum_{I\in B} u_I Y_{[I]},$$

where $e_I = \partial/\partial u_I$.

2.5. The ball-box theorem for free smooth vector fields. We are now ready to draw some consequences from the study of weights of vector fields (Theorem 2.4) in terms of the geometry of balls induced by vector fields. We will get, still in the context of free smooth vector fields, a ball-box theorem which is enough to get a control of the volume of the balls in this setting. In turn, this fact will be exploited in the next subsection to compare the distance induced by lifted vector fields with Rothschild-Stein's quasidistance.

The subelliptic metric introduced by Nagel-Stein-Weinger in [12], in this situation, is defined as follows:

Definition 2.11. For any $\delta > 0$, let $C(\delta)$ be the class of absolutely continuous mappings $\varphi : [0,1] \longrightarrow \Omega$ which satisfy

$$\varphi'(t) = \sum_{|I| \le s} a_I(t)(X_{[I]})_{\varphi(t)} \quad \text{almost everywhere}$$

with

$$|a_I(t)| \le \delta^{|I|}.$$

Then define

$$d(x,y) = \inf\{\delta > 0 : \exists\, \varphi \in C(\delta) \text{ with } \varphi(0) = x,\ \varphi(1) = y\}.$$

Remark 2.12. The quantity $d(x,y)$ is finite for any two points x, $y \in \Omega$. Namely, let $\varphi : [0,1] \longrightarrow \Omega$ be any C^1 curve joining x to y; since the $\{(X_{[I]})_x\}_{|I| \le r}$ span $\mathbb{R}^p$ at any point $x \in \Omega$, $\varphi'(t)$ can always be expressed in the form

$$\sum_{|I| \le r} a_I(t)(X_{[I]})_{\varphi(t)},$$

for suitable bounded functions $a_I(t)$; then the curve φ will belong to the class $C(\delta)$, for $\delta > 0$ large enough, and $d(x,y)$ will be finite and not exceeding this δ.

Proposition 2.13. *The function $d : \Omega \times \Omega \to \mathbb{R}$ is a distance. Moreover, for any $\Omega' \Subset \Omega$ there exist positive constants c_1, c_2 such that*

$$c_1|x - y| \le d(x,y) \le c_2|x - y|^{1/r} \quad \text{for any } x, y \in \Omega'. \tag{2.23}$$

The previous proposition is well known (see [12, Proposition 1.1]; in [2] this is proved also for nonsmooth vector fields).

We are now in position to state our ball-box theorem.

Notation 2.14. For fixed $\bar{x} \in \mathbb{R}^p$, $R > 0$, let

$$\operatorname{Box}(\bar{x}, R) = \Big\{x \in \mathbb{R}^p : x = \exp\Big(\sum_{I \in B} u_I X_{[I]}\Big)(\bar{x}) : |u_I| < R^{|I|} \text{ for any } I \in B\Big\};$$

In canonical coordinates u_I, the subset $\operatorname{Box}(\bar{x}, R)$ simply becomes:

$$\operatorname{Box}(R) = \Big\{u \in \mathbb{R}^p : |u_I| < R^{|I|} \text{ for any } I \in B\Big\}.$$

Let $B(x,R)$ denote the metric ball of center x and radius R in $\mathbb{R}^p$, with respect to the distance d induced by the vector fields $\{X_{[I]}\}_{I \in B}$.

Also, let us define the following quantities related to canonical coordinates:

$$|u|_k = \sum_{|J| = k} |u_J| \quad \text{for } k = 1, 2, \dots, s,$$

$$\|u\| = \sum_{k=1}^{s} |u|_k^{1/k}.$$

Theorem 2.15 (Ball-box theorem for free vector fields). *For every $\Omega' \Subset \Omega$ there exist positive constants C, R_0, c_1, c_2, c_3 depending on Ω, Ω' and the system $\{X_{[I]}\}_{I \in B}$ such that, for any $\bar{x} \in \Omega'$, $R \leq R_0$,*

(i) $\mathrm{Box}(\bar{x}, R) \subseteq B(\bar{x}, R) \subseteq \mathrm{Box}(\bar{x}, CR)$,

(ii) $c_1 R^Q \leq |B(\bar{x}, R)| \leq c_2 R^Q$, *where $Q = \sum_{I \in B} |I|$ plays the role of "homogeneous dimension",*

(iii) $|B(\bar{x}, 2R)| \leq c_3 |B(\bar{x}, R)|$.

Proof. (i) Let us show first that

$$\mathrm{Box}(\bar{x}, R) \subseteq B(\bar{x}, R). \tag{2.24}$$

For $x \in \mathrm{Box}(\bar{x}, R)$, let us write

$$x = \exp\Big(\sum_{I \in B} u_I X_{[I]}\Big)(\bar{x}) \quad \text{with } |u_I| < R^{|I|} \tag{2.25}$$

and set

$$\varphi(t) = \exp\Big(\sum_{I \in B} t u_I X_{[I]}\Big)(\bar{x}).$$

This $\varphi(t)$ defines an admissible curve belonging to $C(\delta)$ for some $\delta < R$, that is, $d(x, \bar{x}) < R$ and inclusion (2.24) is proved.

To prove the reverse inclusion

$$B(\bar{x}, R) \subseteq \mathrm{Box}(\bar{x}, CR),$$

we argue as follows. For $x \in B(\bar{x}, R)$, let $\varphi(t)$ be a curve in $C(R)$, that is,

$$\varphi'(t) = \sum_{I \in B} a_I(t)(X_{[I]})_{\varphi(t)}, \quad \text{with } \varphi(0) = \bar{x},\ \varphi(1) = x,\ |a_I(t)| \leq R^{|I|}.$$

Then, for any smooth function $f(x)$ we have

$$f(\varphi(t)) - f(\bar{x}) = \int_0^t \frac{\mathrm{d}}{\mathrm{d}t}[f(\varphi(\tau))]\,\mathrm{d}\tau = \sum_{I \in B} \int_0^t a_I(\tau)(X_{[I]} f)_{\varphi(\tau)}\,\mathrm{d}\tau.$$

In particular, reasoning from now on in canonical coordinates, for $f(u) = u_J$ we get

$$\varphi(t)_J = \sum_{I \in B} \int_0^t a_I(\tau)(X_{[I]} u_J)_{\varphi(\tau)}\,\mathrm{d}\tau. \tag{2.26}$$

From Theorem 2.4 we read

$$|(X_{[I]} u_J)_{\varphi(\tau)}| \leq c \|\varphi(\tau)\|^{|J| - |I|} \quad \text{if } |I| < |J|,$$

provided x ranges in a compact set. Also, by definition of φ we have

$$|a_I(\tau)| \le R^{|I|} \le CR^{|J|} \quad \text{if } |I| \ge |J|,$$

for any $R \le R_0$, any fixed R_0, and some C depending on R_0. Therefore, (2.26) gives

$$\begin{aligned}(2.27)\quad |\varphi(t)_J|_k &\le C \sum_{I\in B,\, |I|\le k-1} \int_0^t R^{|I|} \|\varphi(\tau)\|^{|J|-|I|}\, \mathrm{d}\tau + \sum_{I\in B,\, |I|\ge k} \int_0^t CR^{|J|}\, \mathrm{d}\tau \\ &= C\Big\{ \sum_{j=1}^{k-1} R^j \int_0^t \|\varphi(\tau)\|^{k-j}\, \mathrm{d}\tau + R^k \Big\} \\ &\le C\Big\{ R \int_0^t \|\varphi(\tau)\|^{k-1}\, \mathrm{d}\tau + R^k \Big\}\end{aligned}$$

where the last inequality holds because for $j = 1, 2, \dots, k-1$

$$R^j \|x\|^{k-j} \le \begin{cases} R^k & \text{if } \|x\| \le R, \\ R\|x\|^{k-1} & \text{if } \|x\| \ge R. \end{cases}$$

Next, since

$$R\|\varphi(\tau)\|^{k-1} \le (R + \|\varphi(\tau)\|)^k \le c(R^k + \|\varphi(\tau)\|^k),$$

from (2.27) we get

$$|\varphi(t)_J|_k \le C\Big\{ \int_0^t \|\varphi(\tau)\|^k\, \mathrm{d}\tau + R^k \Big\}.$$

Recalling the definition of $\|\cdot\|$, we then have

$$\begin{aligned} |\varphi(t)_J|_k^{1/k} &\le C\Big\{ \Big(\int_0^t \|\varphi(\tau)\|^k\, \mathrm{d}\tau)^{1/k} + R \Big\}, \\ \|\varphi(\tau)\| &\le C \sum_{k=1}^{r} \Big\{ \Big(\int_0^t \|\varphi(\tau)\|^k\, \mathrm{d}\tau \Big)^{1/k} + R \Big\}, \\ \|\varphi(\tau)\|^r &\le C \sum_{k=1}^{r} \Big\{ \Big(\int_0^t \|\varphi(\tau)\|^k\, \mathrm{d}\tau \Big)^{r/k} + R^s \Big\} \\ &\qquad\qquad \text{by Hölder's inequality, since } \frac{r}{k} > 1, \\ &\le C \sum_{k=1}^{r} \Big\{ \int_0^t \|\varphi(\tau)\|^r\, \mathrm{d}\tau + R^r \Big\}. \end{aligned}$$

Hence Gronwall's inequality implies

$$\|\varphi(\tau)\|^r \le CR^r \quad \text{for any } \tau < t \le 1,$$

which for $\tau = 1$ gives

$$\|x\| \le CR,$$

that is $x \in \text{Box}(CR)$.

(ii) Let

$$F(u,x) = \exp\Big(\sum_{I\in B} u_I X_{[I]}\Big)(x), \quad \text{for } (u,x) \in U \times \overline{\Omega'}$$

for some neighborhood U of 0, and let $J(u,x)$ be the Jacobian determinant of the map $u \mapsto F(u,x)$. Since

$$J(0,x) = \det((X_{[I]})_x)_{I\in B}$$

by compactness $J(u,x)$ is bounded and bounded away from zero in $U' \times \overline{\Omega'}$, for a suitable open subset $U' \subset U$. Therefore,

$$\begin{aligned} |B(x,R)| &\le |\,\text{Box}(\bar{x},CR)| \\ &= \int_{\text{Box}(\bar{x},CR)} \mathrm{d}y = \int_{|u_I|<(CR)^{|I|}} |J(u,x)|\,\mathrm{d}u \le c\int_{|u_I|<(CR)^{|I|}} \mathrm{d}u = cR^Q, \end{aligned}$$

and analogously we establish the reverse inequality.

Finally, (iii) immediately follows from (ii). □

Remark 2.16. The above proof basically relies on Theorem 2.4 and elementary facts. However, we want also to stress the fact that the uniform control on the constants is possible since the vector fields are free, so that a basis can be chosen once and for all, independently from the point.

2.6. Equivalent quasidistances for free vector fields. Let us consider again the free lifted vector fields $X_0, X_1, \ldots, X_n$, and the map $\Theta(\xi,\eta)$, defined for ξ, η belonging to a suitable neighborhood W of a fixed point ξ_0. Recall that $u = \Theta(\xi,\eta)$ can be seen as an element of the group G. A point $u \in G$ is individuated by coordinates $\{u_J\}_{J\in B}$. One can define

$$\rho(\xi,\eta) = \|\Theta(\xi,\eta)\|,$$

the Rothschild-Stein's *quasidistance* induced by the X_i. It is defined only locally and satisfies the properties collected in the next statement:

Proposition 2.17. *For every $\xi_0 \in \mathbb{R}^p$ there exist a neighborhood W of ξ_0 and constants $c, c_1, c_2 > 0$ such that for any $\xi, \eta, \zeta \in W$,*

$$\rho(\xi,\eta) \geq 0,$$
$$\rho(\xi,\eta) = 0 \iff \xi = \eta,$$
$$\rho(\xi,\eta) = \rho(\eta,\xi),$$
$$\rho(\xi,\eta) \leq c\{\rho(\xi,\zeta) + \rho(\zeta,\eta)\}, \tag{2.28}$$
$$c_1 d(\xi,\eta) \leq \rho(\xi,\eta) \leq c_2 d(\xi,\eta). \tag{2.29}$$

Proof. The first three properties follow by definition and by Proposition 2.6, while (2.28) follows by (2.29), since d satisfies the triangle inequality. So let us prove (2.29).

Let us denote by $B_\rho(\xi,R)$ the "balls" with respect to ρ. Note that, just by definition of Box and Θ, we have

$$\eta \in \mathrm{Box}(\xi,R) \iff \|\Theta(\xi,\eta)\| \leq R,$$

which implies the inclusions

$$B_\rho(\xi,c_1R) \subset \mathrm{Box}(\xi,R) \subset B_\rho(\xi,c_2R)$$

for any $\xi \in W$, $R \leq R_0$, and some positive constants R_0, c_1, c_2. Therefore, Theorem 2.15 implies (2.29). ∎

We also have the following result.

Proposition 2.18. *The change of coordinate in $\mathbb{R}^p$ given by*

$$\xi \mapsto u = \Theta(\xi,\eta)$$

has a Jacobian determinant given by

$$\mathrm{d}\xi = c(\eta)(1 + O(\|u\|))\,\mathrm{d}u$$

where $c(\eta)$ is a smooth function, bounded and bounded away from zero.

The above proposition is proved in [14]; see also [1, Theorem 1.7]. Moreover, this proposition is a particular case of the analogous property which we will prove for nonsmooth vector fields in Subsection 3.4, Proposition 3.11.

3. APPROXIMATION FOR NONSMOOTH HÖRMANDER'S VECTOR FIELDS

Here we want to prove also for nonsmooth vector fields the approximation theorem and the basic results about the map $\Theta_\eta(\cdot)$. This is made possible combining the previous theory for smooth vector fields with a natural procedure of approximation of nonsmooth vector fields by their Taylor expansion. To quantify the weight of the "remainders" in the approximation formula, we have to assume the coefficients of the vector fields in the scale of Hölder spaces, slightly strengthening the assumptions made in Section 1.2.

3.1. Assumptions.

Assumptions (B). We assume that for some integer $r \geq 2$, some $\alpha \in (0,1]$ and some bounded domain $\Omega \subset \mathbb{R}^p$ the following hold:

(B1) The coefficients of the vector fields $X_1, X_2, \ldots, X_n$ belong to $C^{r-1,\alpha}(\Omega)$, while the coefficients of X_0 belong to $C^{r-2,\alpha}(\Omega)$. Here and in the following, $C^{k,\alpha}$ stands for the classical space of functions with derivatives up to order k, Hölder continuous of exponent α.

(B2) The vectors $\{(X_{[I]})_x\}_{|I|\leq r}$ span $\mathbb{R}^p$ at every point $x \in \Omega$.

The following easy lemma is proved analogously to that in Section 1.1.

Lemma 3.1. *Under the assumption* (B1) *above, for any* $1 \leq k \leq r$*, the differential operators*

$$\{X_I\}_{|I|\leq k}$$

are well defined, and have $C^{r-k,\alpha}$ *coefficients. The same is true for the vector fields* $\{X_{[I]}\}_{|I|\leq k}$.

Dependence of the constants We will often write that some constant depends on the vector fields X_i's and some fixed domain $\Omega' \Subset \Omega$. (Actually, the dependence on the X_i's will be usually left understood). Explicitly, this will mean that the constant depends on:

(i) Ω';

(ii) The norms $C^{r-1,\alpha}(\Omega)$ of the coefficients of X_i $(i = 1, 2, \ldots, n)$ and the norms $C^{r-2,\alpha}(\Omega)$ of the coefficients of X_0;

(iii) A positive constant c_0 such that the following bound holds:

$$\inf_{x\in\Omega'} \max_{|I_1|,|I_2|,\ldots,|I_p|\leq r} \left| \det((X_{[I_1]})_x, (X_{[I_2]})_x, \ldots, (X_{[I_p]})_x) \right| \geq c_0$$

(where "det" denotes the determinant of the $p \times p$ matrix having the vectors $(X_{[I_i]})_x$ as rows).

Note that (iii) is a quantitative way of assuring the validity of Hörmander's condition, uniformly in Ω'.

As we have seen in Section 1.2, we can always lift our nonsmooth vector fields to get a system of free vector fields, still satisfying assumptions (B1), (B2). In this section we are interested in proving a Rothschild-Stein type approximation result for these lifted vector fields. Therefore, just to simplify notation, throughout this section we will assume that our vector fields X_i's are already free up to weight r.

3.2. Regularized canonical coordinates. As we have seen in Section 2.1, of basic importance in the study of the properties of the vector fields is expressing them in terms of *canonical coordinates*. This means to introduce the local diffeomorphism

$$\mathbb{R}^p \ni u \longmapsto x = \exp\Big(\sum_{I\in B} u_I X_{[I]}\Big)(\bar{x}) \in U(\bar{x})$$

and then express the vector fields $X_{[I]}^x$ as differential operators $X_{[I]}^u$ acting on the variable u. However, under our assumptions, we cannot expect this map being more regular than Hölder continuous; even if we strengthened our assumptions asking the coefficients of X_i to be $C^{r-p_i+1,\alpha}$, we would get a $C^{1,\alpha}$ local diffeomorphism, which would transform the vector fields X_i^x in C^α vector fields X_i^u. Therefore it would be impossible to compute the commutators of the transformed vector fields, and all the arguments of Section 2.1 would break down. More precisely, following this line we would be forced to require the coefficients of X_i to be $C^{2r,\alpha}$, which is rather unsatisfactory. This discussion leads us to look for a smooth diffeomorphism, adapted to the system $\{X_{[I]}(x)\}_{|I|\le r}$, transforming the vector fields $X_{[I]}$ in vector fields $X_{[I]}^u$ having the same regularity, and better properties. This leads to the concept of *regularized canonical coordinates*, firstly introduced in [6] and then used by several authors in particular cases.

Fix a point $\bar{x} \in \Omega$; for any $i = 0, 1, 2, \dots, n$, let us consider the vector field

$$X_i = \sum_{j=1}^{p} b_{ij}(x)\,\partial_{x_j};$$

let $p_{ij}^r(x)$ be the Taylor polynomial of $b_{ij}(x)$ of center $\bar{x}$ and order $r - p_i$; note that

$$b_{ij}(x) = p_{ij}^r(x) + O(|x-\bar{x}|^{r-p_i+\alpha}); \tag{3.1}$$

set

$$S_i^{\bar{x}} = \sum_{j=1}^{p} p_{ij}^r(x)\,\partial_{x_j}.$$

From (3.1) we easily have (see, e.g., [2]) the following result:

Proposition 3.2. *The* $S_i^{\bar{x}}$ $(i = 0, 1, 2, \dots, n)$ *are smooth vector fields defined in the whole space, satisfying:*

$$(S_I^{\bar{x}})_{\bar{x}} = (X_I)_{\bar{x}} \text{ and } (S_{[I]}^{\bar{x}})_{\bar{x}} = (X_{[I]})_{\bar{x}} \quad \text{for any } I \text{ with } |I| \le r.$$

Moreover,

$$X_{[I]} - S_{[I]}^{\bar{x}} = \sum_{j=1}^{p} c_I^j(x)\,\partial_{x_j} \quad \text{with } c_I^j(x) = O(|x-\bar{x}|^{r-|I|+\alpha}).$$

Finally, denoting by d_X and $d_{S^{\bar{x}}}$ the distances (see Section 2.5) induced by the X_i's and the $S_i^{\bar{x}}$'s, respectively, and by B_X and $B_{S^{\bar{x}}}$ the corresponding metric balls, there exist positive constants c_1, c_2, R_0 depending on Ω, Ω' and the X_i's, but not on $\bar{x} \in \Omega'$, such that

$$B_{S^{\bar{x}}}(\bar{x}, c_1 R) \subset B_X(\bar{x}, R) \subset B_{S^{\bar{x}}}(\bar{x}, c_2 R)$$

for any $R < R_0$.

Now, fix a point $\bar{x} \in \Omega$, and select a basis of $\mathbb{R}^p$ of the form

$$\{(X_{[I]})_{\bar{x}}\}_{I \in B}.$$

Clearly, we have

$$\{(X_{[I]})_{\bar{x}}\}_{I \in B} = \{(S_{[I]})_{\bar{x}}\}_{I \in B}.$$

We can now introduce, as in Section 2.1, the canonical coordinates induced by the smooth vector fields $S_i^{\bar{x}}$; these will be, by definition, the regularized canonical coordinates induced by the X_i's:

$$\mathbb{R}^p \ni u \longmapsto x = \exp\Big(\sum_{I \in B} u_I S_{[I]}^{\bar{x}}\Big)(\bar{x}) \tag{3.2}$$

for x belonging to some neighborhood $U(\bar{x})$. Note that the Jacobian of the map $u \longmapsto x$, at $u = 0$, equals the matrix of the vector fields $\{(S_{[I]}^{\bar{x}})_{\bar{x}}\}_{I \in B} = \{(X_{[I]})_{\bar{x}}\}_{I \in B}$, therefore is nonsingular. Moreover, since the $S_{[I]}^{\bar{x}}$'s are smooth, the diffeomorphism is smooth, too.

Next, we express our original vector fields X_i in terms of regularized canonical coordinates u: let us write $X_{[I]}^u$ to denote the vector field $X_{[I]}$ expressed in coordinates u. The following facts are immediate:

Proposition 3.3.

(i) *The (transformed) vector field X_i^u has $C^{r-p_i,\alpha}$ coefficients, for $i = 0, 1, 2, \dots, n$;*

(ii) *The vector field $X_{[I]}^u$ has $C^{r-|I|,\alpha}$ coefficients, for any I such that $|I| \le r$; in particular, all the $\{X_{[I]}^u\}_{I \in B}$ have $C^{0,\alpha}$ coefficients;*

(iii) *We have*

$$\Big[X_{[I]}, X_{[J]}\Big]^u = \Big[X_{[I]}^u, X_{[J]}^u\Big]$$

for any I, J such that $|I| + |J| \le r$.

3.3. Weights and approximation. Using the coordinates u we can give the following definition.

Definition 3.4. Let X be a vector field with possibly nonsmooth coefficients. We will say that X has weight $\ge k \in \mathbb{R}$ near the point $\bar{x}$ if, expressing it in regularized canonical coordinates

$$X^u = \sum_{J \in B} c_J(u)\, \partial_{u_J},$$

we have:

$$|c_J(u)| \le c\|u\|^{k+|J|}.$$

for u in a neighborhood of 0.

Here, as in Section 2.1,

$$\|u\| = \sum_{J\in B} |u_J|^{1/|J|}.$$

Note that if X is a smooth vector field of weight $\ge k \in \mathbb{Z}$, in the sense of Definition 2.2, then it is also of weight $\ge k$ in the sense of the above definition. We will also use the following elementary remark:

> If X, Y have weight $\ge k_X, k_Y$, respectively, then $X \pm Y$ has weight $\ge \min(k_X, k_Y)$.

Proposition 3.5. *The vector field $X_{[I]} - S^{\bar{x}}_{[I]}$ has weight $\ge \alpha - |I|$ near $\bar{x}$, for any $|I| \le r$.*

Proof. By Proposition 3.2, we know that

$$X_{[I]} - S^{\bar{x}}_{[I]} = \sum_{j=1}^{p} c_I^j(x)\,\partial_{x_j} \quad \text{with } c_I^j(x) = O(|x-\bar{x}|^{r-|I|+\alpha}). \tag{3.3}$$

Let

$$u = F(x) = (F_J(x))_{J\in B}$$

be the local smooth diffeomorphism defined as in (3.2), and let $x = F^{-1}(u)$ be its inverse. Then vector fields are transformed according to the law:

$$\partial_{x_j} = \sum_{J\in B} (\partial_{x_j} F_J)(x)\,\partial_{u_J}.$$

Therefore

$$X^u_{[I]} - (S^{\bar{x}}_{[I]})^u = \sum_{J\in B}\sum_{j=1}^{p} (c_I^j\,\partial_{x_j} F_J)(x)\,\partial_{u_J} = \sum_{J\in B} \tilde{c}_I^J(u)\,\partial_{u_J}$$

with

$$|\tilde{c}_I^J(u)| = \Big|\sum_{j=1}^{p} (c_I^j\,\partial_{x_j} F_J)(x)\Big| \quad \text{since } F \text{ is a smooth diffeomorphism and by (3.3)}$$

$$\le c\sum_{j=1}^{p} |c_I^j(x)| \le c|x-\bar{x}|^{r-|I|+\alpha} \le c\|u\|^{r-|I|+\alpha} \le c\|u\|^{|J|-|I|+\alpha}.$$

This ends the proof. □

Proposition 3.6. *For any $|I| \le r$ we have:*

(i) *The vector field $X_{[I]}$ has weight $\ge -|I|$ near $\bar{x}$.*
(ii) *If $Y_0, Y_1, \dots, Y_n$ is any system of smooth vector fields satisfying (with respect to regularized canonical coordinates)*

$$\sum_{I \in B} u_I e_I = \sum_{I \in B} u_I Y_{[I]},$$

then $X_{[I]} - Y_{[I]}$ has weight $\ge \alpha - |I|$.

Proof. We can apply to the system of smooth vector fields $\{S_{[I]}^{\bar{x}}\}_{I \in B}$ the theory developed in Section 2 and say that:

(i) The vector field $S_{[I]}^{\bar{x}}$ has weight $\ge -|I|$.
(ii) $S_{[I]}^{\bar{x}} - Y_{[I]}$ has weight $\ge 1 - |I|$.

Assertion (ii) exploits the fact that the $S_i^{\bar{x}}$'s are free up to weight r at $\bar{x}$, if the X_i's are so, because the X_i's and the $S_i^{\bar{x}}$'s satisfy the same commutation relations, up to weight r, at $\bar{x}$.

Therefore, by Proposition 3.5, we conclude that:

(i) $X_{[I]} = S_{[I]}^{\bar{x}} + (X_{[I]} - S_{[I]}^{\bar{x}})$ has weight $\ge \min(-|I|, -|I| + \alpha) = -|I|$,
(ii) $X_{[I]} - Y_{[I]} = (X_{[I]} - S_{[I]}^{\bar{x}}) + (S_{[I]}^{\bar{x}} - Y_{[I]})$
has weight $\ge \min(\alpha - |I|,\ 1 - |I|) = \alpha - |I|$. □

We can now deduce from the previous proposition a local approximation result, analogous to Theorem 2.7:

Theorem 3.7. *If $X_{[I]}^u$ denotes the vector field $X_{[I]}$ expressed in regularized canonical coordinates centered at $\bar{x}$, and $Y_{[I]}$ is a left invariant homogeneous vector field on the group G, as above, then*

$$X_{[I]}^u = Y_{[I]} + R_{\bar{x},[I]},$$

where $R_{\bar{x},[I]}$ is a $C^{r-|I|,\alpha}$ vector field of weight $\ge \alpha - |I|$ near $\bar{x}$, depending on $\bar{x}$ in a C^α continuous way.

Proof. Fix a point $\bar{x} \in \mathbb{R}^p$, and define the smooth approximating vector fields $S_{i,\bar{x}}$. Here it will be more convenient to denote by the subscript $\bar{x}$ the dependence on the center $\bar{x}$ of the approximation. We know that, expressing the vector fields with respect to the canonical coordinates u of $S_{[I],\bar{x}}$ at $\bar{x}$ (regularized canonical coordinates of $X_{[I]}$ at $\bar{x}$),

$$X_{[I]}^u - S_{[I],\bar{x}}^u \quad \text{has weight } \ge \alpha - |I| \text{ near } u = 0, \text{ for any } |I| \le r.$$

More explicitly, this means that we can write

$$X^u_{[I]} = S^u_{[I],\bar{x}} + O_{[I],\bar{x}}, \tag{3.4}$$

where $O_{[I],\bar{x}}$ are $C^{r-|I|,\alpha}$ vector fields (in the variables u) of weight $\geq \alpha - |I|$ near $u = 0$, and their coefficients are C^α functions of $\bar{x}$, because the same is true for $S^u_{[I],\bar{x}}$. This last assertion follows directly by the definition of $S_{i,\bar{x}}$ and our assumptions, in view of the following remark:

$$\text{If } X_i = \sum_{j=1}^{p} b_{ij}(x)\,\partial_{x_j}, \quad \text{then } S_{i,\bar{x}} = \sum_{j=1}^{p} \Big(\sum_{|\alpha| \leq r - p_i} \frac{\partial_x^\alpha b_{ij}(\bar{x})}{\alpha!} (x - \bar{x})^\alpha \Big)\,\partial_{x_j}.$$

Since we are assuming $b_{ij} \in C^{r-p_i,\alpha}$, from the above formula one reads that:

(i) The coefficients of $S_{i,\bar{x}}$ are C^α functions of $\bar{x}$;
(ii) The same is true for commutators $S_{[I],\bar{x}}$ for $|I| \leq r$;
(iii) The same is true if we express $S_{[I],\bar{x}}$ with respect to new variables u which are smooth functions of x.

This completes the proof of C^α dependence of $O_{[I],\bar{x}}$ on $\bar{x}$.

We now consider the $S_{i,\bar{x}}$'s as smooth vector fields defined in the whole space $\mathbb{R}^p$ and, for any fixed $\eta \in \mathbb{R}^p$, we apply Rothschild-Stein's local approximation theorem to the smooth $S_{i,\bar{x}}$'s, writing

$$S^v_{[I],\bar{x}} = Y_{[I]} + \hat{R}^{\bar{x}}_{\eta,[I]}, \tag{3.5}$$

where $Y_{[I]}$ are left invariant vector fields on the group, $\hat{R}^{\bar{x}}_{\eta,[I]}$ are smooth vector fields of weight $\geq 1 - |I|$, smoothly depending on the point η, and the superscript v in $S^v_{[I],\bar{x}}$ means that these vector fields are expressed with respect to the canonical coordinates v of $S_{[I],\bar{x}}$ centered at η. The vector fields $\hat{R}^{\bar{x}}_{\eta,[I]}$ also depend on $\bar{x}$, because a different $\bar{x}$ means a different set of vector fields $S^v_{[I],\bar{x}}$. Since, by point (iii) here above, $S^v_{[I],\bar{x}}$ depend on $\bar{x}$ in a C^α-continuous way, the same is true for $\hat{R}^{\bar{x}}_{\eta,[I]}$.

Next, we set $\eta = \bar{x}$ in (3.5); then $v = u$ (canonical coordinates of $S_{[I],\bar{x}}$ centered at $\bar{x}$), so we can write

$$S^u_{[I],\bar{x}} = Y_{[I]} + \hat{R}^{\bar{x}}_{\bar{x},[I]}$$

where $\hat{R}^{\bar{x}}_{\bar{x},[I]}$ is a smooth vector field of weight $\geq 1 - |I|$ near $\bar{x}$, depending on $\bar{x}$ in a C^α continuous way. This fact, together with (3.4), allows us to write:

$$X^u_{[I]} = Y_{[I]} + R_{\bar{x},[I]}$$

where $R_{\bar{x},[I]}$ is a $C^{r-|I|,\alpha}$ vector field of weight $\geq \alpha - |I|$ near $\bar{x}$, depending on $\bar{x}$ in a C^α continuous way. □

To make more usable the previous theorem we have to construct, as in the smooth case, a map $u = \Theta(\xi_0, \xi)$, allowing to compute the derivative $X_{[I]}f(\xi)$ without passing to variables u.

Let:

$$\xi = E(u, \xi_0) = \exp\Big(\sum_{I \in B} u_I S_{[I],\xi_0}\Big)(\xi_0). \tag{3.6}$$

Clearly, for any fixed ξ_0, the map $u \longmapsto E(u, \xi_0)$ is smooth. Moreover, its Jacobian determinant at $u = 0$ equals

$$\det(S_{[I],\xi_0})(\xi_0) \neq 0$$

because $\{(S_{[I],\xi_0})_{\xi_0}\}_{I \in B}$ is a basis of $\mathbb{R}^p$. Therefore there exists a smooth inverse function, which we denote by

$$u = \Theta_{\xi_0}(\xi).$$

A basic difference with the smooth theory is that $\Theta_{\xi_0}(\xi)$ is not simply $-\Theta_{\xi}(\xi_0)$. This is due to the fact that if $\xi = E(u, \xi_0)$, then

$$E(-u, \xi) = \exp\Big(-\sum_{I \in B} u_I S_{[I],\xi}\Big) \exp\Big(\sum_{I \in B} u_I S_{[I],\xi_0}\Big)(\xi_0) \neq \xi_0$$

because the vector fields in the first exponential are $S_{[I],\xi}$ while those in the second one are $S_{[I],\xi_0}$. Due to this asymmetry, we cannot expect $\Theta_{\xi_0}(\xi)$ to be as smooth in the ξ_0 variable as it is in the ξ variable. Instead, since the vector fields depend on ξ_0 in a C^α continuous way, the best we can hope is C^α continuity with respect to ξ_0 also for $\Theta_{\xi_0}(\xi)$. This is actually the case, as shown below:

Proposition 3.8. *For any fixed ξ, the map $\xi_0 \mapsto \Theta_{\xi_0}(\xi)$ is C^α.*

Proof. We start noting that the function $\xi_0 \mapsto E(u, \xi_0)$ is C^α continuous. This follows from (3.6) by continuous dependence estimates on the exponential, keeping in mind that $\xi_0 \mapsto S_{[I],\xi_0}$ is C^α.

We are now going to revise the proof of the inverse function theorem, showing that in this case a C^α dependence on the parameter ξ_0 of the function $u \mapsto E(u, \xi_0)$ implies a C^α dependence on the same parameter ξ_0 for the inverse function $\xi \mapsto \Theta_{\xi_0}(\xi)$.

Let A_{ξ_0} be the Jacobian matrix $\partial_u E(u, \xi_0)$ evaluated at $u = 0$, and, for a fixed ξ, set

$$\varphi_{\xi_0}(u) = u + A_{\xi_0}^{-1}(E(u, \xi_0) - \xi).$$

To find the inverse function of $u \mapsto E(u, \xi_0)$ we look for a fixed point of φ_{ξ_0}. Since $E(u, \xi_0)$ is a smooth function of u and $E(0, \xi_0) - \xi_0 = 0$, for u in a suitably small neighborhood of 0 and ξ in a suitable neighborhood of ξ_0, φ_{ξ_0} is a contraction; under these assumptions, we can write

$$|\varphi_{\xi_0}(u_1) - \varphi_{\xi_0}(u_2)| \le \delta|u_1 - u_2|$$

for some $\delta \in (0,1)$. For any two points ξ_1, ξ_2 in a small neighborhood of ξ_0, let us define the sequence:

$$\begin{cases} u_{n+1}^{\xi_i} = \varphi_{\xi_i}(u_n), \\ u_0^{\xi i} = 0, \end{cases}$$

for $i = 1, 2$. Clearly, $u_n^{\xi_i} \to E(\cdot, \xi_i)^{-1}(\xi) \equiv u^{\xi_i}$ and

$$\begin{aligned} u_{n+1}^{\xi_1} - u_{n+1}^{\xi_2} &= u_n^{\xi_1} - u_n^{\xi_2} + A_{\xi_1}^{-1}(E(u_n^{\xi_1}, \xi_1) - \xi) - A_{\xi_2}^{-1}(E(u_n^{\xi_2}, \xi_2) - \xi) \\ &= \{u_n^{\xi_1} - u_n^{\xi_2} + A_{\xi_1}^{-1}(E(u_n^{\xi_1}, \xi_1) - E(u_n^{\xi_2}, \xi_1))\} \\ &\quad + \{(A_{\xi_2}^{-1} - A_{\xi_1}^{-1})(\xi)\} + \{A_{\xi_1}^{-1}(E(u_n^{\xi_2}, \xi_1) - E(u_n^{\xi_2}, \xi_2))\} \\ &\quad + \{(A_{\xi_1}^{-1} - A_{\xi_2}^{-1})E(u_n^{\xi_2}, \xi_2)\} \\ &\equiv \{A\} + \{B\} + \{C\} + \{D\}. \end{aligned}$$

Now,

$$|\{A\}| = |\varphi_{\xi_1}(u_n^{\xi_1}) - \varphi_{\xi_1}(u_n^{\xi_2})| \le \delta|u_n^{\xi_1} - u_n^{\xi_2}|$$

while

$$|\{B\}| + |\{C\}| + |\{D\}| \le c|\xi_1 - \xi_2|^\alpha$$

(where we used the fact that $E(u, \xi_0)$ is C^α in ξ_0, uniformly in u, for small u). Hence

$$|u_{n+1}^{\xi_1} - u_{n+1}^{\xi_2}| \le \delta|u_n^{\xi_1} - u_n^{\xi_2}| + c|\xi_1 - \xi_2|^\alpha.$$

Passing to the limit for $n \to \infty$ we get

$$|u^{\xi_1} - u^{\xi_2}| \le \delta|u^{\xi_1} - u^{\xi_2}| + c|\xi_1 - \xi_2|^\alpha$$

and so

$$|u^{\xi_1} - u^{\xi_2}| \le \frac{c}{1-\delta}|\xi_1 - \xi_2|^\alpha$$

which is the desired C^α continuous dependence. □

Finally, by Theorem 3.7 and Proposition 3.8 we immediately get the following result:

Theorem 3.9. *Let $Y_{[I]}$ be the left invariant homogeneous vector field on the group G. Then*

$$X^{\xi}_{[I]}(f(\Theta_\eta(\xi))) = (Y_{[I]}f + R_{\eta,[I]}f)(\Theta_\eta(\xi)), \tag{3.7}$$

where $\Theta_\eta(\cdot)$ is a smooth diffeomorphism, depending on η in a C^α continuous way, and $R_{\eta,[I]}$ are $C^{r-|I|,\alpha}$ vector fields of weight $\geq \alpha - |I|$, depending on η in a C^α continuous way.

3.4. Equivalent quasidistances and properties of the nonsmooth map Θ. Here we will prove two useful properties of the map $\Theta_\eta(\cdot)$.

Proposition 3.10. *For every $\xi_0 \in \mathbb{R}^p$ there exist a neighborhood W of ξ_0 and constants C_1, $C_2 > 0$ such that for any ξ, $\eta \in W$, the following local equivalence holds:*

$$C_1 d(\xi,\eta) \leq \rho(\xi,\eta) \leq C_2 d(\xi,\eta).$$

Proof. By the ball-box theorem for free smooth vector fields, we have

$$\rho(\xi,\eta) = \|\Theta_\eta(\xi)\| \simeq d_{\tilde{S}\eta}(\eta,\xi).$$

In turn, by Proposition 3.2,

$$d_{\tilde{S}\eta}(\eta,\xi) \simeq d_{\tilde{X}}(\eta,\xi)$$

and we are done. ◻

Proposition 3.11. *The change of coordinate in $\mathbb{R}^p$ given by*

$$u = \Theta_\eta(\xi)$$

has a Jacobian determinant given by

$$\mathrm{d}\xi = c(\eta)(1 + O(\|u\|))\,\mathrm{d}u$$

where $c(\eta)$ is a C^α function, bounded and bounded away from zero. More explicitly, this means that $\mathrm{d}\xi = [c(\eta) + \omega(\eta,u)]\,\mathrm{d}u$ with $|\omega(\eta,u)| \leq c\|u\|$ and $c(\eta)$ as above.

Proof. We will compute the Jacobian determinant of the inverse mapping $\xi = \Theta_\eta^{-1}(u)$. To do this, set

$$X_{[I]} = \sum_{k=1}^{N} c_{Ik}(\xi)\frac{\partial}{\partial \xi_k} \quad \text{for every } I \in B$$

and rewrite the left hand side of (3.7) as

$$\sum_k c_{Ik}(\xi)\sum_j \frac{\partial f}{\partial u_j}(\Theta_\eta(\xi))\frac{\partial}{\partial \eta_k}\left[(\Theta_\eta(\xi))_j\right].$$

Then (3.7), evaluated at $\xi = \eta$, becomes:

$$\sum_k c_{Ik}(\eta)\sum_j \frac{\partial f}{\partial u_j}(0)\frac{\partial}{\partial \eta_k}\left[(\Theta_\eta(\xi))_j\right]_{\xi=\eta} = \left(Y_{[I]}f + R^\eta_{[I]}f\right)(0).$$

Choosing $f(u) = u_J$ $(J \in B)$,

$$\sum_k c_{Ik}(\eta)\frac{\partial}{\partial \eta_k}\left[(\Theta_\eta(\xi))_J\right]_{\eta=\xi} = \left(Y_{[I]}u_J + R^\eta_{[I]}u_J\right)(0) = \delta_{IJ}$$

where the last equality follows recalling that

$$Y_{[I]}[f](0) = \frac{\mathrm{d}}{\mathrm{d}t} f(\exp tY_{[I]})_{/t=0}$$

and that $\exp tY_{[I]}$ equals, in local coordinates, $(0,\dots,t,\dots,0)$ with t in the $[I]$-th position. As to $R^\eta_{[I]}u_J$, by Theorem 3.9 it has weight $\geq \alpha - |I|$, which by definition means that

$$R^\eta_{[I]} = \sum_{J\in B} a_{IJ}(u)\frac{\partial}{\partial u_J} \quad \text{with } |a_{IJ}(u)| \leq C\|u\|^{\alpha-|I|+|J|}.$$

Then

$$R^\eta_{[I]}u_J = \delta_{IJ}a_{IJ}(u);$$
$$|R^\eta_{[I]}u_J| \leq C\|u\|^\alpha;$$
$$(R^\eta_{[I]}u_J)(0) = 0.$$

Defining the square matrix

$$C(\eta) = \{c_{hk}(\eta)\}_{hk}$$

and letting $J(\eta)$ be the Jacobian determinant of the mapping $u = (\Theta_\eta(\xi))$ at $\xi = \eta$, we get

$$\mathrm{Det}[C(\eta)]\cdot J(\eta) = 1.$$

Hence the Jacobian determinant of the mapping $\xi = \Theta_\eta^{-1}(u)$ at $u = 0$ equals $\mathrm{Det}[C(\eta)] \equiv c(\eta)$, which is a C^α function, as the coefficients of the vector fields

$X_{[I]}$ are. Moreover, $c(\eta)$ is bounded away from zero since the $X_{[I]}$'s are independent.

Since the determinant of $\xi = \Theta_\eta^{-1}(u)$ is a smooth function in u, it equals

$$c(\eta) + \omega(\eta, u)$$

with $|\omega(\eta, u)| \le c\|u\|$ and we conclude

$$\mathrm{d}\xi = c(\eta) \cdot (1 + O(\|u\|))\,\mathrm{d}u. \qquad \square$$

Acknowledgement. We wish to thank Professor A. Melin, who kindly accepted to clarify to us some arguments contained in his paper [11].

REFERENCES

[1] M. BRAMANTI and L. BRANDOLINI, *L^p estimates for nonvariational hypoelliptic operators with VMO coefficients*, Trans. Amer. Math. Soc. **352** (2000), no. 2, 781–822. http://dx.doi.org/10.1090/S0002-9947-99-02318-1. MR1608289 (2000c:35026)

[2] M. BRAMANTI, L. BRANDOLINI, and M. PEDRONI, *Basic properties of nonsmooth Hörmander's vector fields and Poincaré's inequality*, Forum Mathematicum, to appear, available at http://dx.doi.org/10.1515/FORM.2011.133.

[3] M. BRAMANTI, L. BRANDOLINI, M. MANFREDINI, and M. PEDRONI, *Fundamental solutions and local solvability for nonsmooth Hörmander's operators*, preprint.

[4] A. BONFIGLIOLI and F. UGUZZONI, *Families of diffeomorphic sub-Laplacians and free Carnot groups*, Forum Math. **16** (2004), no. 3, 403–415. http://dx.doi.org/10.1515/form.2004.018. MR2050190 (2005f:22029)

[5] M. CHRIST, A. NAGEL, E.M. STEIN, and S. WAINGER, *Singular and maximal Radon transforms: analysis and geometry*, Ann. of Math. (2) **150** (1999), no. 2, 489–577. http://dx.doi.org/10.2307/121088. MR1726701 (2000j:42023)

[6] G. CITTI, *C^∞ regularity of solutions of a quasilinear equation related to the Levi operator*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **23** (1996), no. 3, 483–529. MR1440031 (98b:35072)

[7] G.B. FOLLAND, *Subelliptic estimates and function spaces on nilpotent Lie groups*, Ark. Mat. **13** (1975), no. 1–2, 161–207. http://dx.doi.org/10.1007/BF02386204. MR0494315 (58 #13215)

[8] ———, *On the Rothschild-Stein lifting theorem*, Comm. Partial Differential Equations **2** (1977), no. 2, 165–191. http://dx.doi.org/10.1080/03605307708820028. MR0433514 (55 #6490)

[9] R. GOODMAN, *Lifting vector fields to nilpotent Lie groups*, J. Math. Pures Appl. (9) **57** (1978), no. 1, 77–85. MR0494316 (58 #13216)

[10] L. HÖRMANDER, *Hypoelliptic second order differential equations*, Acta Math. **119** (1967), no. 1, 147–171. http://dx.doi.org/10.1007/BF02392081. MR0222474 (36 #5526)

[11] L. HÖRMANDER and A. MELIN, *Free systems of vector fields*, Ark. Mat. **16** (1978), no. 1, 83–88. http://dx.doi.org/10.1007/BF02385983. MR0650825 (58 #31284)

[12] A. NAGEL, E.M. STEIN, and S. WAINGER, *Balls and metrics defined by vector fields. I. Basic properties*, Acta Math. **155** (1985), no. 1-2, 103–147. http://dx.doi.org/10.1007/BF02392539. MR793239 (86k:46049)

[13] F. RICCI, *Sub-Laplacians on Nilpotent Lie Groups: Course Notes*, available at http://homepage.sns.it/fricci/papers/sublaplaciani.pdf.

[14] L. P. ROTHSCHILD and E.M. STEIN, *Hypoelliptic differential operators and nilpotent groups*, Acta Math. **137** (1976), no. 3–4, 247–320. http://dx.doi.org/10.1007/BF02392419. MR0436223 (55 #9171)

MARCO BRAMANTI:
Dipartimento di Matematica
Politecnico di Milano
Via Bonardi 9
20133 Milano, Italy
E-MAIL: marco.bramanti@polimi.it

LUCA BRANDOLINI & MARCO PEDRONI:
Dipartimento di Ingegneria dell'Informazione e Metodi Matematici
Università di Bergamo
Viale Marconi 5
24044 Dalmine BG, Italy
E-MAIL: luca.brandolini@unibg.it
E-MAIL: marco.pedroni@unibg.it

KEY WORDS AND PHRASES: nonsmooth Hörmander's vector fields, lifting, subelliptic distance.

2000 MATHEMATICS SUBJECT CLASSIFICATION: 53C17.

Received: *February 5th, 2010.*

Article electronically published on June 19th, 2010.