Rev. Mat. Iberoamericana **27** (2011), no. 2, 645–732

Multi-parameter Carnot-Carathéodory balls and the theorem of Frobenius

Brian Street

Abstract

We study multi-parameter Carnot-Carathéodory balls, generalizing results due to Nagel, Stein and Wainger in the single parameter setting. The main technical result is seen as a uniform version of the theorem of Frobenius. In addition, we study maximal functions associated to certain multi-parameter families of Carnot-Carathéodory balls.

1. Introduction

In the seminal paper [22], Nagel, Stein and Wainger gave a detailed study of Carnot-Carathéodory balls. The main purpose of this paper is to develop an analogous theory of *multi-parameter* Carnot-Carathéodory balls: a situation where the methods of [22] do not apply in general. We will see that the main results for multi-parameter Carnot-Carathéodory balls follow from a certain "uniform" version of the theorem of Frobenius on involutive distributions.[1] We will prove this version of the theorem of Frobenius by building on the work of [22] along with work of Tao and Wright [30]. Our primary motivation is to obtain the properties of multi-parameter balls which are relevant for developing a theory of multi-parameter singular integrals, which will be the subject of a future paper. To this end, we will estimate the volume of certain multi-parameter balls, and we will study maximal functions associated to certain families of multi-parameter balls. In addition, we will study the composition of certain "unit operators."

2000 Mathematics Subject Classification: Primary: 53C17; Secondary: 53C12, 42B25.
Keywords: Sub-Riemannian geometry, Carnot-Carathéodory geometry, Frobenius theorem, multi-parameter, spaces of homogeneous type, maximal functions.

[1] Here, and in the rest of the paper, we are considering (possibly) singular distributions. That is, the dimension of the distribution may vary from point to point.

We begin by introducing the notion of a Carnot-Carathéodory ball. Suppose we are given q C^1 vector fields, $X_1, \ldots, X_q$ on an open set $\Omega \subseteq \mathbb{R}^n$; denote this list of vector fields by X. We define the Carnot-Carathéodory ball of unit radius, centered at $x_0 \in \Omega$, with respect to the list X by:[2]

$$B_X(x_0) := \Bigg\{ y \in \Omega \,\Bigg|\, \exists \gamma : [0,1] \to \Omega, \gamma(0) = x_0, \gamma(1) = y,$$
$$\gamma'(t) = \sum_{j=1}^{q} a_j(t) X_j(\gamma(t)), a_j \in L^\infty([0,1]), \left\| \Big(\sum_{1 \leq j \leq q} |a_j|^2 \Big)^{\frac{1}{2}} \right\|_{L^\infty([0,1])} < 1 \Bigg\}.$$

Now that we have the definition for Carnot-Carathéodory balls with unit radius, we may define Carnot-Carathéodory balls of any radius merely by scaling the vector fields. This leads us directly to multi-parameter balls, of which the single parameter balls of [22] are a special case.

Fix $\nu \geq 1$, an integer. We will discuss ν-parameter balls. To each vector field X_j, we associate a formal degree $0 \neq d_j \in [0, \infty)^\nu$. We denote by (X, d) the list of vector fields

$$(X_1, d_1), \ldots, (X_q, d_q).$$

Furthermore, for $\delta \in [0, \infty)^\nu$, we denote by $\delta^d X$ the list of vector fields:

$$\delta^{d_1} X_1, \ldots, \delta^{d_q} X_q$$

where δ^{d_j} is defined by the standard multi-index notation. That is, $\delta^{d_j} = \prod_{\mu=1}^{\nu} \delta_\mu^{d_j^\mu}$. Then we define the multi-parameter Carnot-Carathéodory ball centered at $x_0 \in \Omega$ of radius δ by:

$$B_{(X,d)}(x_0, \delta) := B_{\delta^d X}(x_0).$$

The theory in [22] concerns the case when $\nu = 1$ (see Section 1.2.1 for a discussion of their results). One of the main goals of this paper is to develop appropriate conditions on the list (X, d) to allow for a general theory of such multi-parameter balls.

It has long been understood that singular integrals corresponding to the single parameter balls of [22] play a fundamental role in many questions in the regularity of linear partial differential operators that are defined by vector fields; in particular, they arise in many questions in several complex variables. This began in [6, 24], and was followed by [25, 7, 12]. These works were followed by many others; too many to offer a detailed account here.

[2]Here, and in the rest of the paper, we write $\gamma'(t) = Z(t)$ to mean $\gamma(t) = \gamma(0) + \int_0^t Z(s)\, ds$.

In the area of several complex variables, some examples are [3, 18, 4, 13]. Recently, however, *multi-parameter* singular integrals, where the underlying geometries are non-Euclidean, have been shown to arise in various special cases in several complex variables and in the parametricies for certain linear partial differential operators. Moreover, these examples are not even amenable to the usual product theory of singular integrals (as is covered in, for example, [20]): the geometries overlap in a non-trivial way. In this vein see [16, 17, 21, 28]. It is our hope that this paper will help play a role in unlocking more general theories.

1.1. Informal statement of results and outline of the paper

In this section, we offer a brief overview of some of the key results of the paper. One of the main aspects of the proofs, and of the interrelationships between the results, is keeping careful track of parameters the constants in the results depend on. This makes the rigorous formulation of these results somewhat technical. Because of this, in this section, we state the results only in the C^∞ category (while we will later deal with less smoothness) and are not precise about what parameters the constants depend on. After each result we will refer the reader to the part of the paper which contains the precise formulation of the result. In addition, Theorem 1.2 represents only a special case of the main result of the paper (Theorem 5.3).

Before we begin, we need a few pieces of notation. Given two integers $1 \leq m \leq n$, we let $\mathcal{I}(m,n)$ be the set of all lists of integers $(i_1,\ldots,i_m)$, such that:

$$1 \leq i_1 < i_2 < \cdots < i_m \leq n.$$

Furthermore, suppose A is an $n \times q$ matrix, and suppose $1 \leq n_0 \leq n \wedge q$, for $I \in \mathcal{I}(n_0,n)$, $J \in \mathcal{I}(n_0,q)$ define the $n_0 \times n_0$ matrix $A_{I,J}$ by using the rows from A which are listed in I and the columns of A which are listed in J. We define:

$$\det_{n_0\times n_0} A = \left(\det A_{I,J}\right)_{\substack{I\in\mathcal{I}(n_0,n)\\ J\in\mathcal{I}(n_0,q)}}.$$

In particular, $\det_{n_0\times n_0} A$ is a *vector*. It will not be important to us in which order the coordinates are arranged. For further information on this object, see Appendix B.

For a vector $v \in \mathbb{R}^n$, we write $|v|$ for the usual ℓ^2 norm, and $|v|_\infty$ and $|v|_1$ for the ℓ^∞ and ℓ^1 norms, respectively. For a matrix A, we write $\|A\|$ for the usual operator norm. Finally, we write $B_n(\eta)$ for the ball in $\mathbb{R}^n$, centered at 0, of radius $\eta > 0$ in the $|\cdot|$ norm.

The setup of the main result is as follows. We are given q C^∞ vector fields $X_1,\ldots,X_q$ defined on a fixed open set $\Omega \subseteq \mathbb{R}^n$. Corresponding to each vector field we are given a formal degree $0 \neq d_j \in [0,\infty)^\nu$, where ν is a fixed positive integer. We let (X,d) denote the list of vector fields with formal degrees $(X_1,d_1),\ldots,(X_q,d_q)$, and we let X denote the list of vector fields $X_1,\ldots,X_q$. At times, we will identify X with the $n\times q$ matrix whose columns are given by $X_1,\ldots,X_q$ (similarly for other lists of vector fields). Our main assumption is that for every $\delta\in[0,1)^\nu$, with $|\delta|$ sufficiently small,[3] we have:

$$\left[\delta^{d_j}X_j,\delta^{d_k}X_k\right]=\sum_{l=1}^{q} c_{j,k}^{l,\delta}\delta^{d_l}X_l. \tag{1.1}$$

We assume that $c_{j,k}^{l,\delta}\in C^\infty$ *uniformly* in δ; i.e., that as δ varies, $c_{j,k}^{l,\delta}$ varies over a bounded subset of C^∞.[4]

Remark 1.1. Note that we have not assumed that the list of vector fields X spans the tangent space. This will prove the be an essential point in much of what follows. One thing to observe is that while X may not span the tangent space, (1.1) implies that the distribution spanned by $\delta^d X$ is involutive, and therefore the classical theorem of Frobenius applies to show that these vector fields foliate Ω into leaves (see Section 1.2.5 for a review of the classical theorem of Frobenius). The Carnot-Carathéodory ball $B_{(X,d)}(x_0,\delta)$ is then an open subset of the leaf passing through x_0 generated by this distribution. In what follows, we will estimate the volume of this ball (denoted by $\mathrm{Vol}\left(B_{(X,d)}(x_0,\delta)\right)$). This volume is taken in the sense of the induced Lebesgue measure on the leaf.

For $n_0\leq q$ and $J=(j_1,\ldots,j_{n_0})\in\mathcal{I}(n_0,q)$, we write $(X,d)_J$ to denote the list of vector fields with formal degrees $(X_{j_1},d_{j_1}),\ldots,\left(X_{j_{n_0}},d_{j_{n_0}}\right)$, and we write X_J to denote the list of vector fields $X_{j_1},\ldots,X_{j_{n_0}}$, similarly we write d_J for the list of formal degrees $d_{j_1},\ldots,d_{j_{n_0}}$. For each $x\in\Omega$, let $n_0(x,\delta)=\dim\operatorname{span}\left\{\delta^{d_1}X_1(x),\ldots,\delta^{d_q}X_q(x)\right\}$,[5] and for each $x\in\Omega$, and δ sufficiently small pick $J(x,\delta)\in\mathcal{I}(n_0(x,\delta),q)$ such that:

$$\left|\det_{n_0(x,\delta)\times n_0(x,\delta)}\delta^{d_{J(x,\delta)}}X_{J(x,\delta)}(x)\right|_\infty=\left|\det_{n_0(x,\delta)\times n_0(x,\delta)}\delta^{d}X(x)\right|_\infty.$$

[3] Throughout the rest of this introduction, δ will always denote a small element of $[0,1)^\nu$.

[4] Even in the smooth case, the assumptions in Section 5.1 require less than we outline here.

[5] Note, the dependence of $n_0(x,\delta)$ on δ only involves which of the coordinates of δ are 0.

For $u \in \mathbb{R}^{n_0(x,\delta)}$ with $|u|$ sufficiently small, define the map[6]

$$\Phi_{x,\delta}(u) = e^{u\cdot\left(\delta^d X\right)_{J(x,\delta)}}x = e^{u\cdot \delta^{d_{J(x,\delta)}} X_{J(x,\delta)}}x.$$

Our main theorem is:

Theorem 1.2. *Let K be a compact subset of Ω. Then, there exist constants $\eta, \xi \approx 1$ such that for all δ sufficiently small and all $x \in K$:*

$$B_{(X,d)}(x, \xi\delta) \subseteq \Phi_{x,\delta}\left(B_{n_0(x,\delta)}(\eta)\right) \subseteq B_{(X,d)}(x,\delta)$$

and

1. $\Phi_{x,\delta} : B_{n_0(x,\delta)}(\eta) \to B_{(X,d)}(x_0, \delta)$ *is one-to-one.*
2. *For all* $u \in B_{n_0(x,\delta)}(\eta)$,
$$\left| \det_{n_0(x,\delta)\times n_0(x,\delta)} d\Phi_{x,\delta}(u)\right| \approx \left|\det_{n_0(x,\delta)\times n_0(x,\delta)} \delta^d X(x)\right|.$$
3. $\mathrm{Vol}\left(B_{(X,d)}(x,\delta)\right) \approx \left|\det_{n_0(x,\delta)\times n_0(x,\delta)} \delta^d X(x)\right|$. *This is essentially a consequence of Items* 1 and 2.
4. $\mathrm{Vol}\left(B_{(X,d)}(x,2\delta)\right) \lesssim \mathrm{Vol}\left(B_{(X,d)}(x,\delta)\right)$. This is essentially a consequence of Item 3.

In addition to what is stated in Theorem 1.2, a number of other technical results hold which are essential for applications. In particular, the map $\Phi_{x,\delta}$ can be used as a "scaling" map. This is because the pullback of the vector fields $\delta^d X$ via the map $\Phi_{x,\delta}$ to $B_{n_0(x,\delta)}(\eta)$ satisfy good properties uniformly in x and δ.[7] We refer the reader to Section 5 for a discussion of these results along with the rigorous statement of Theorem 1.2. Note that in the single parameter case, Item 4 is the main inequality that must be satisfied for the balls $B_{(X,d)}(x,\delta)$ to form a space of homogeneous type (when paired with Lebesgue measure). This is the first sign that these multi-parameter balls, in this generality, will yield analogs to some results from the single-parameter Calderón-Zygmund theory.

As was mentioned earlier, Theorem 1.2 will follow from a "uniform" version of the theorem of Frobenius. To understand this connection, one must

[6] If Z is a C^1 vector field, then $e^Z x$ is defined in the following way. Let $E(t)$ be the unique solution to the ODE $\frac{d}{dt}E(t) = Z(E(t))$, $E(0) = x$. Then, $e^Z x$ is defined to be $E(1)$, provided this solution exists up to $t = 1$ (which it will if Z has sufficiently small C^1 norm). See Appendix A for further details.

[7] See Section 5.2.4 to see a scaling technique in action.

first understand the connection between multi-parameter balls and single parameter balls. Given the multi-parameter formal degrees $0 \neq d_j \in [0,\infty)^\nu$, we obtain corresponding single parameter degrees, which we denote by $\sum d$, and are defined by $\left(\sum d\right)_j := \sum_{\mu=1}^{\nu} d_j^\mu = |d_j|_1$. Given δ, we decompose $\delta = \delta_0\delta_1$, where $\delta_0 \in [0,\infty)$ and $\delta_1 \in [0,\infty)^\nu$.[8] Then, directly from the definition, we obtain:

$$B_{(X,d)}\left(x_0,\delta\right) = B_{\left(\delta_1^d X,\sum d\right)}\left(x_0,\delta_0\right) = B_{\left(\delta^d X,\sum d\right)}\left(x_0,1\right).$$

Because of this, to prove Theorem 1.2 for a fixed $x \in K$ and a fixed δ, it suffices to prove a result for a list of vector fields with single-parameter formal degrees: the vector fields $\left(\delta^d X, \sum d\right)$.

At this point, we change notation. We now work in the single-parameter case $\nu = 1$. We suppose we are given q C^∞ vector fields $X_1,\ldots,X_q$ on a fixed open set $\Omega \subseteq \mathbb{R}^n$ and associated to each X_j we are given a formal degree $d_j \in (0,\infty)$. We further suppose that we are given a fixed point $x_0 \in \Omega$.[9] One should think of this single-parameter list (X,d) as coming from a multi-parameter list via $\left(\delta^d X, \sum d\right)$ as in Theorem 1.2. Our main assumption is that we have:

$$[X_j, X_k] = \sum_{l=1}^{q} c_{j,k}^l X_l \tag{1.2}$$

where $c_{j,k}^l \in C^\infty$.

Let $n_0 = \dim \operatorname{span}\{X_1(x_0),\ldots,X_q(x_0)\}$, and pick $J \in \mathcal{I}(n_0,q)$ such that:

$$\left|\det_{n_0\times n_0} X_J(x_0)\right|_\infty = \left|\det_{n_0\times n_0} X(x_0)\right|_\infty.$$

For $u \in \mathbb{R}^{n_0}$ with $|u|$ sufficiently small, define the map

$$\Phi(u) = e^{u\cdot X_J}x_0.$$

In what follows, the constants can be chosen uniformly as the X_j and $c_{j,k}^l$ vary over bounded subsets of C^∞. The constants do *not* depend on a lower bound for, say, $|\det_{n_0\times n_0} X(x_0)|$. Our "uniform" version of the theorem of Frobenius is:

Theorem 1.3. *There exist $\eta,\xi \approx 1$ such that:*

$$B_{(X,d)}(x_0,\xi) \subseteq \Phi\left(B_{n_0}(\eta)\right) \subseteq B_{(X,d)}(x_0,1)$$

[8] Of course this decomposition is not unique.

[9] In addition, we need to assume that x_0 is not too close to the boundary of Ω, but we ignore such technicalities in this introduction.

and

- $\Phi : B_{n_0}(\eta) \to B_{(X,d)}(x_0, 1)$ *is one-to-one.*
- *For all* $u \in B_{n_0}(\eta)$, $|\det_{n_0 \times n_0} d\Phi(u)| \approx |\det_{n_0 \times n_0} X(x_0)|$.

Furthermore, if we let Y_j *be the pullback of* X_j *via the map* Φ*, then the list of vector fields* $Y_1, \ldots, Y_q$ *satisfy good estimates. See Theorem* 4.1 *for details.*

Let us now describe why Theorem 1.3 can be viewed as a version of the theorem of Frobenius (see Section 1.2.5 for further discussion on this point). Indeed, our main assumption (1.2) is exactly the main assumption of the theorem of Frobenius. Hence under the hypotheses of Theorem 1.3, the vector fields $X_1, \ldots, X_q$ foliate Ω into leaves. As mentioned in Remark 1.1, $B_{(X,d)}(x_0, \xi)$ is an open neighborhood of x_0 on this leaf. Moreover $\Phi : B_{n_0}(\eta) \to B_{(X,d)}(x_0, 1)$ is one-to-one. Thus, Φ can be considered as a coordinate chart on the leaf in a neighborhood of x_0. Hence for each point $x_0 \in \Omega$, Theorem 1.3 yields a coordinate chart near x_0 on the leaf passing through x_0. In this way, Theorem 1.3 *implies* the classical theorem of Frobenius. The main point is that not only does Theorem 1.3 yield a coordinate chart, but it also allows one to take $\xi, \eta \approx 1$ and it gives good estimates on this coordinate chart; estimates which do not follow from the standard proofs of the theorem of Frobenius (see Remark 3.4), nor from the methods of [22] (see the discussion in Section 1.2.1).

In Section 1.2, we discuss a number of previous, related works, and relate our results to these works. In Section 3 we state and prove a precise version of Theorem 1.3 in the special case when $X_1(x_0), \ldots, X_q(x_0)$ are linearly independent, and it is in this section that the main technicalities of the paper lie. We refer to this result as a uniform theorem of Frobenius. The proof heavily uses methods from Section 4 of [30] and methods from [22], but these need to be significantly generalized to adapt them to our situation. In Section 4 we use the results of Section 3 to prove the more general version of Theorem 1.3 in the case when $X_1(x_0), \ldots, X_q(x_0)$ are not necessarily linearly independent. We refer to this as studying Carnot-Carathéodory balls "at the unit scale."[10] In addition, we use these results to define smooth bump functions supported on these balls, along the lines of those used in [19]. In Section 5 we state and prove the rigorous version of Theorem 1.2. In Section 6, we use the results from Section 4.2 to study multi-parameter

[10]Here we mean at the unit scale with respect to the vector fields X_j. Thus, if the X_j are very small (as is the case when $X_j = \delta^{d_j} W_j$, where δ is small), then one can think of it as being at a small scale. In addition, we could have equally well referred to this as a theorem of Frobenius. We chose this name, though, to emphasize its role in the proof of Theorem 1.2.

maximal functions associated to a certain subclass of our multi-parameter balls. In Section 6 we also discuss compositions of certain unit operators (see Corollary 6.8 and Section 1.2.4), and in Section 6.1 we use these unit operators to discuss the relationship between certain quasi-metrics that arise.

From the discussion proceeding Theorem 1.3, it is clear why Theorem 1.3 implies Theorem 1.2, provided one has appropriate control over the implicit constants in Theorem 1.3. Hence a main aspect of this paper is to keep track of the appropriate constants in Theorem 1.3. At times, this will be quite technical. In addition, we will state our main results with only a finite amount of smoothness, further complicating our notations.[11] In the past, there has been some interest in results in the single parameter case, using as low regularity as possible. Even in this single parameter context, our results are new in this direction. See Section 1.2.3 for a discussion of this.

In an effort to ease the notation in the paper, at the start of many of the sections of this paper, we will define a notion of "admissible constants." These will be constants that only depend on certain parameters. This notion of admissible constant may change from section to section, but we will be explicit about what it means each time. In addition, if κ is another parameter, and we say "there exists an admissible constant $C = C(\kappa)$," we mean that C is allowed to depend on everything an admissible constant may depend on, and is also allowed to depend on κ. We use the notation $A \lesssim B$ to mean $A \leq CB$, where C is an admissible constant; so that, in particular, the meaning of $\lesssim$ may change from section to section. We use $A \approx B$ to mean $A \lesssim B$ and $B \lesssim A$. In some sections, we will use different levels of smoothness assumptions. In these sections, we will also define a notion of m-admissible constants, where $m \in \mathbb{N}$ denotes the level of assumed smoothness. We will write $A \lesssim_m B$ for $A \leq CB$, where C is an m-admissible constant, and we define $\approx_m$ in a similar manner.

We write $Q_n(\eta)$ to denote the unit ball in $\mathbb{R}^n$, centered at 0, of radius η in the $|\cdot|_\infty$ norm. All functions in this paper are assumed to be real valued. Given a, possibly not closed, set $U \subseteq \mathbb{R}^n$, we write:

$$\|f\|_{C^m(U)} = \sup_{x \in U} \sum_{|\alpha| \leq m} |\partial_x^\alpha f(x)|.$$

Finally, $v_1, v_2 \in \mathbb{R}^m$ are two vectors, we write $v_1 \leq v_2$ to mean that the inequality holds for each coordinate.

Remark 1.4. Throughout the paper we work on an open subset $\Omega \subseteq \mathbb{R}^n$, endowed with Lebesgue measure. At first glance, it might seem useful to work more generally on a Riemannian manifold (where Lebesgue measure is repla-

[11] While it does complicate notation, working with only a explicit finite amount of smoothness does not complicate our proof.

ced by the volume element); and replace the set of vector fields $X_1, \dots, X_q$ with a locally finitely generated distribution (endowed with an appropriate (multi-parameter) filtration taking the place of the formal degrees). However, our results are local in nature, and working in such a setting offers no new generality and only serves to complicate notation.

1.2. Past work

In this section we discuss other results from the literature which are related to the results in this paper. In particular, we discuss the work of [22] and the work in Section 4 of [30]. Next, we discuss other results concerning Carnot-Carathéodory balls in the case when the vector fields are not smooth. In particular, we discuss the recent works [1] and [15]. Third, as motivation for our study of "unit operators" (and maximal functions) in Section 6, we discuss the singular integrals from [28]. Finally, we discuss the classical theorem of Frobenius and make some further remarks on how Theorem 1.3 can be seen as a "uniform" version.

1.2.1. Single-parameter balls and the work of Nagel, Stein, and Wainger

In this section, we discuss the main results of [22]. In fact, their main results can be seen as a special case of Theorem 1.2, in the single-parameter case ($\nu = 1$).

We are given an open set $\Omega \subseteq \mathbb{R}^n$ and C^∞ vector fields $X_1, \dots, X_q$ on Ω, with corresponding formal degrees $d_1, \dots, d_q \in (0, \infty)$. [22] assumes two properties of the vector fields and formal degrees:

1. There exist $c_{i,j}^k \in C^\infty$ such that

$$[X_i, X_j] = \sum_{d_k \leq d_i + d_j} c_{i,j}^k X_k. \tag{1.3}$$

2. The vector fields $X_1, \dots, X_q$ span the tangent space at every point.

In this context, Nagel, Stein, and Wainger prove Theorem 1.2 (for a fixed compact set $K \Subset \Omega$). Note that Item 1 is a special case of (1.1). Indeed, one may take

$$c_{i,j}^{k,\delta} = \begin{cases} \delta^{d_i + d_j - d_k} c_{i,j}^k & \text{if } d_k \leq d_i + d_j, \\ 0 & \text{otherwise.} \end{cases}$$

The implicit constants are allowed to depend not only on upper bounds for a finite number of the C^m norms of the X_j and the $c_{i,j}^k$ (as in Theorem 1.2), but also a lower bound for:

$$\inf_{x\in K}\left|\det_{n\times n} X(x)\right|. \tag{1.4}$$

This is the fundamental difference between the results of [22] and Theorem 1.2.

Indeed, using the connection between single-parameter and multi-parameter balls discussed in Section 1.1, it is not hard to see that Theorem 1.2 is essentially *equivalent* to obtaining the results of [22] without allowing the constants to depend on a lower bound for (1.4). Of course, if one does not allow the constants to depend on a lower bound for (1.4), one should also consider the limiting result when the quantity in (1.4) is equals 0. I.e., when the vector fields do not span the tangent space at every point. This is precisely the statement of Theorem 1.2 in the single parameter case.

Use of a lower bound for (1.4) is essential to the methods of [22]. It is used, for instance, every time the error term in the Campbell-Hausdorff formula is estimated.[12] To explain this, we outline a proof of (a result similar to) Lemma 2.13 of [22]. We take the setting as above, and consider the map ($B_q(\eta)\to\Omega$, for some small $\eta>0$):

$$\theta_\delta(s) = e^{s_1\delta^{d_1}X_1+\ldots+s_q\delta^{d_q}X_q}x_0.$$

Then one has:

$$d\theta_\delta\left(\partial_{s_j}\right) = \sum_{j=1}^{q} c_j^{k,\delta}\delta^{d_k}X_k,$$

with $c_j^{k,\delta}$ bounded uniformly for $\delta>0$ small. Indeed, the Campbell-Hausdorff formula allows one to compute the Taylor series for $d\theta_\delta\left(\partial_{s_j}\right)$. One has, for every $N>0$,

$$\begin{aligned} d\theta_\delta\left(\partial_{s_j}\right) = {}& \delta^{d_j}X_j + a_1\left[s\cdot\delta^d X, \delta^{d_j}X_j\right] + a_2\left[s\cdot\delta^d X,\left[s\cdot\delta^d X,\delta^{d_j}X_j\right]\right]+\cdots \\ & + a_{N-1}\left\{\text{commutators of order } N-1\right\} + O\big(\left|\delta^d s\right|^N\big), \end{aligned}$$

where the a_j are constants and $\delta^d s = \left(\delta^{d_1}s_1,\ldots,\delta^{d_q}s_q\right)$. The first N terms are of the desired form by (1.3) (or more generally, (1.1)). Thus, the goal is to see that $O\big(\left|\delta^d s\right|^N\big)$ is of the desired form. This can be seen directly, by taking N so large that $N\min_j\{d_j\}\geq\max_j\{d_j\}$, and using the lower bound for (1.4). However, this procedure does not work in the multi-parameter situation. Indeed, consider the two-parameter situation. In the case when $\delta_1 << \delta_2$, then the best one can say about the error term $O\big(\left|\delta^d s\right|^N\big)$ is that it

[12] See the appendix of [22] for an introduction to the Campbell-Hausdorff formula.

is bounded by as large a power of δ_2 as we like (by taking N large). However, we would need it to be bounded by a large power of δ_1 to generalize the above proof. It turns out that, even in the multi-parameter situation, the error term *is* of the desired form. This follows *a fortiori* from the results of this paper. Because of this, one can use the results of this paper to apply the proofs in [22] to the multi-parameter situation. However, since the results in [22] follow from the results in this paper, this idea does not improve the main results of this paper. This idea does have some uses, though: one can often "lift" results from the single-parameter setting to the multi-parameter setting by using the results from this paper. This is discussed in more detail in Section 5.2.4.

At first glance, one might think that the proper generalization of (1.3) to the multi-parameter situation would be:

$$[X_i, X_j] = \sum_{d_k \leq d_i + d_j} c_{i,j}^k X_k, \tag{1.5}$$

where $d_j \in [0,\infty)^\nu$ and the inequality $d_k \leq d_i + d_j$ is meant coordinatewise. Just as before this is a special case of (1.1), and one may take

$$c_{i,j}^{k,\delta} = \begin{cases} \delta^{d_i+d_j-d_k} c_{i,j}^k & \text{if } d_k \leq d_i + d_j, \\ 0 & \text{otherwise.} \end{cases}$$

However, unlike in the single-parameter case, (1.5) does not encapsulate a large fraction of the interesting examples. This is explained in more detail in Section 5.3.

1.2.2. Weakly comparable balls, the work of Tao and Wright, and a motivating example

In this section, we discuss the work in Section 4 of [30] on "weakly-comparable" Carnot-Carathéodory balls. While the results discussed in this section do not follow from Theorem 1.2, they do follow from the more general Theorem 5.3–this is discussed in Section 5.2.1.

To understand these results, we must first understand the main motivating example of [22]. Suppose we are given C^∞ vector fields $W_1, \ldots, W_r$ on an open subset $\Omega \subseteq \mathbb{R}^n$. Suppose further that these vector fields satisfy Hörmander's condition: i.e., $W_1, \ldots, W_r$ along with their commutators up to some fixed finite order (say, up to order $m \in \mathbb{N}$) span the tangent space at every point. We assign to each vector field W_j the formal degree 1. We assign to each commutator $[W_i, W_j]$ the formal degree 2. We assign to each commutator $[W_i, [W_j, W_k]]$ the formal degree 3. We continue this process up to degree m, and we obtain a list of vector fields with one parameter formal degrees $(X_1, d_1), \ldots, (X_q, d_q)$. As usual, we denote this list by (X, d). It is easy to check that this list of vector fields satisfies the assumptions in Section 1.2.1. It is also shown in [22] that the single-parameter balls

$$B_{\delta W_1, \ldots, \delta W_r}(x_0)$$

are comparable to the single-parameter balls

$$B_{(X,d)}(x_0, \delta).$$

Because of this, one can use the results of [22] to study the balls

$$B_{\delta W_1, \ldots, \delta W_r}(x_0).$$

We now turn to discussing two-parameter weakly-comparable balls. The restriction to two-parameters is not essential, see Section 5.2.1. We suppose again that we are given a list of C^∞ vector fields satisfying Hörmander's condition, $W_1, \ldots, W_r$. We now separate this list into two lists:

$$W_1', \ldots, W_{r_1}', \quad W_1'', \ldots, W_{r_2}'',$$

so that the two lists together satisfy Hörmander's condition, but they may not satisfy Hörmander's condition separately. Suppose we wish to study the two-parameter balls given by

$$B_{\delta_1 W', \delta_2 W''}(x_0),$$

where $\delta_1, \delta_2 \in (0, 1)$ are small. Thus, when $\delta_1 = \delta_2$ this reduces to the single-parameter case discussed above. It is natural to wish for an estimate of the form

$$\operatorname{Vol}(B_{2\delta_1 W', 2\delta_2 W''}(x_0)) \lesssim \operatorname{Vol}(B_{\delta_1 W', \delta_2 W''}(x_0)). \tag{1.6}$$

Unfortunately, (1.6) does not hold in general (See Example 5.6). To obtain (1.6), there are three options:

1. We could restrict the vector fields we consider. This is the perspective taken up in Theorem 1.2.
2. We could restrict the form of $\delta = (\delta_1, \delta_2)$. This is the perspective taken up in Section 4 of [30].
3. We could, more generally, do a combination of the above two methods. This is taken up in Theorem 5.3.

We briefly discuss the first two methods, and refer the reader to Theorem 5.3 for the third.

The first method is quite straight forward given the Theorem 1.2. We assign to each of the vector fields $W_1', \dots, W_{r_1}'$ the formal degree $(1,0)$. We assign to each of the vector fields $W_1'', \dots, W_{r_1}''$ the formal degree $(0,1)$. If we have assigned a vector field Z_1 the formal degree d_1 and Z_2 the formal degree d_2, we assign to the commutator $[Z_1, Z_2]$ the formal degree $d_1 + d_2$. One then uses this procedure to take commutators of the W' and W'' to some large finite order, thereby yielding a list of vector fields with two-parameter formal degrees $(X_1, d_1), \dots, (X_q, d_q)$. The restriction we put on the vector fields W' and W'' is merely that this list of vector fields satisfies the conditions of Theorem 1.2.[13] One can then apply Theorem 1.2 to study these balls.

For the second method, Tao and Wright noted that one does not have to restrict the vector fields one considers, provided one restricts attention to δ which are "weakly-comparable". To define this notion, fix large constants κ, N. We then restrict our attention to $\delta = (\delta_1, \delta_2)$ that satisfy:

$$\frac{1}{\kappa}\delta_1^N \leq \delta_2 \leq \kappa \delta_1^{\frac{1}{N}}.$$

In this case, one can prove develop a very satisfactory theory of the balls with these radii. In particular, one has (1.6). See Section 5.2.1 for more details.

Despite the fact, as is mentioned in [30], that the proofs from [22] generalize to show everything they needed, Tao and Wright put forth another proof method. That these methods can be rephrased, generalized, and combined with the methods of [22] and classical methods to prove more general results is one of the main points of this paper.

Remark 1.5. One point we have skipped over in this section is to compare the balls $B_{\delta W}(x_0)$ to the balls $B_{(X,d)}(x_0, \delta)$ in the multi-parameter situation (they turn out to be comparable). This is taken up in Section 5.2.4.

1.2.3. Nonsmooth Hörmander vector fields

The results in [22] were stated only for C^∞ vector fields. However, it is clear from their work that there is a finite number M (depending on various quantities) such that one need only consider C^M vector fields. Unfortunately, this M is quite large. This is due to the fact that the uses of the

[13]It is a consequence of the results in Section 5.3 that the balls one obtains in this manner are essentially independent of the order of commutators one takes. That is, if the vector fields (X, d) were obtained by taking commutators up to order M, and if (X, d) satisfies the assumptions of Theorem 1.2, then the vector fields obtained by taking commutators up to order $M + 1$ also satisfy the assumptions of Theorem 1.2, and yield comparable balls.

Campbell-Hausdorff formula in [22] require using very high order Taylor approximations of many of the functions involved.

Much work has been done, in this single-parameter setting, to reduce the required regularity in the results of [22]. Quite recently, and independently of this paper, two works have made great strides on this problem: the work of Barmanti, Brandolini, and Pedroni [1] and the work of Montanari and Morbidelli [15]. We refer the reader to these to works for the long list of works that proceeded them and for a description of applications for such results.

To describe these results, suppose we are given vector fields $W_1, \ldots, W_r$ satisfying Hörmander's condition at step $s \geq 1$. That is, $W_1, \ldots, W_r$ along with their commutators up to order s span the tangent space at each point (see the discussion at the start of Section 1.2.2). Then, [1] shows that one can recreate much of the theory of [22] provided one assumes the vector fields are C^{s-1}. [15] achieves the same thing assuming the vector fields lie in a space that is between $C^{s-2,1}$ and $C^{s-1,1}$ (see [15] for a precise statement).

The regularity assumptions in this paper are incomparable to those discussed above.[14] As far as *isotropic* estimates go, the work of [1, 15] requires less regularity than ours. However, our estimates are non-isotropic in nature. To understand this, use the vector fields $W_1, \ldots, W_r$ to generate a list of vector fields with single-parameter formal degrees $(X_1, d_1), \ldots, (X_q, d_q)$ as in the start of Section 1.2.2. We then assume that each X_j is C^2, and assume a non-isotropic estimate on the $c_{i,j}^k$ which is weaker than assuming the $c_{i,j}^k$ are in C^2 (here, the $c_{i,j}^k$ are as in (1.3)). Note that if one were to replace this with an isotropic estimate, we would require that W_j be in C^{s+2}, which is much worse than the results in [1, 15]. However, the point here is that we do not need to take derivatives of W_j in *every* direction up to order $s+2$, but instead we can mostly restrict our attention to derivatives that arise from taking commutators.

It is likely that the regularity required in this paper is not minimal–even for the methods we use. Indeed, we often show that a subset of C^1 is precompact by showing that it is bounded in C^2, leaving much room for improvement. Improving this would require an even more detailed study of the various ODEs that arise than is already undertaken in this paper, and this would take us quite afield of the main purpose of this paper (to understand the multi-parameter situation). Ideally, one would like to unify the non-isotropic estimates in this paper with the isotropic estimates of [1, 15], we do not attempt to do so here but we hope that the results in this paper will help to motivate future work in this direction.

---

[14] Of course, our results also apply to the multi-parameter situation, which is not true of [1, 15].

1.2.4. Some multi-parameter singular integrals

In [28], an algebra of singular integral operators was developed which contained both the left and right invariant Calderón-Zygmund operators on a stratified Lie group (see [28] for a precise statement). In this section, we outline the main technical estimate that was key to the work of [28]. One of the main results of Section 6 is a generalization[15] of this estimate, and we refer to the operators that come into play in this estimate as "unit operators."

For the purposes of this section, we discuss only the case of the three dimensional Heisenberg group $\mathbb{H}^1$, though all of the results discussed here hold more generally on stratified Lie groups. As a manifold $\mathbb{H}^1$ is diffeomorphic to $\mathbb{R}^3$. For an introduction to $\mathbb{H}^1$, the reader may consult Chapter XII of [27]. If we write $(x, y, t) \in \mathbb{R}^3$ for coordinates on $\mathbb{H}^1$, then the group law is given by:

$$(x, y, t)\,(x', y', t') = (x + x', y + y', t + t' + 2\,(yx' - xy'))\,.$$

With this group law, $\mathbb{H}^1$ is a three dimensional, nilpotent Lie group. As such, it has a three dimensional Lie algebra. The left invariant vector fields are spanned by:

$$X_L = \partial_x + 2y\partial_t, \quad Y_L = \partial_y - 2x\partial_t, \quad T = \partial_t.$$

The right invariant vector fields are spanned by:

$$X_R = \partial_x - 2y\partial_t, \quad Y_R = \partial_y + 2x\partial_t, \quad T = \partial_t.$$

Note that we have:

$$[X_L, Y_L] = -4T, \quad [X_R, Y_R] = 4T,$$

and T commutes with all of the vector fields. Moreover, the left invariant vector fields commute with the right invariant vector fields.

Using these vector fields, we may create a two-parameter list of vector fields given by:

$$(X_L, (1, 0))\,, (Y_L, (1, 0))\,, (T, (2, 0))\,, (X_R, (0, 1))\,, (Y_R, (0, 1))\,, (T, (0, 2))\,,$$

and we denote this list by (X, d). It is easy to see that (X, d) satisfies the assumptions of Theorem 1.2; we therefore obtain a theory of the two-parameter Carnot-Carathéodory balls $B_{(X,d)}\,(x_0, \delta)$.

Let χ be the characteristic function of the unit ball in $\mathbb{R}^3$, and for $r \in (0, \infty)$, define

$$\chi_r\,(x, y, t) = r^4 \chi\left(rx, ry, r^2 t\right).$$

It is easy to see that $\chi_r(\xi^{-1}\zeta)$ is supported for ξ essentially[16] in $B_{(X,d)}\big(\zeta, \big(\tfrac{1}{r}, 0\big)\big)$.

[15]See Corollary 6.8 for the statement of this generalization.

[16]By this we mean it is supported in a comparable ball.

Moreover, it is bounded by a constant times $\mathrm{Vol}\left(B_{(X,d)}\left(\zeta,\left(\frac{1}{r},0\right)\right)\right)^{-1}$ (and these results are sharp). For $\chi_r\left(\zeta\xi^{-1}\right)$, the same is true, but one must use the radius $\left(0,\frac{1}{r}\right)$ instead of $\left(\frac{1}{r},0\right)$.

Define a left invariant operator and a right invariant operator by:

$$\mathrm{Op}_L\left(\chi_r\right): f\mapsto f*\chi_r, \quad \mathrm{Op}_R\left(\chi_r\right): f\mapsto \chi_r * f.$$

A key ingredient of the theory in [28] was a study of the Schwartz kernel of the operator $\mathrm{Op}_L\left(\chi_{r_1}\right)\mathrm{Op}_R\left(\chi_{r_2}\right)$ (see Section 5.1 of [28]). Let $K_{r_1,r_2}\left(\zeta,\xi\right)$ denote this Schwartz kernel. The results in [28] show:

- $K_{r_1,r_2}\left(\zeta,\xi\right)$ is supported essentially in $B_{(X,d)}\left(\zeta,\left(\frac{1}{r_1},\frac{1}{r_2}\right)\right)$.
- $K_{r_1,r_2}\left(\zeta,\xi\right)\lesssim \mathrm{Vol}\left(B_{(X,d)}\left(\zeta,\left(\frac{1}{r_1},\frac{1}{r_2}\right)\right)\right)^{-1}$.
- The above two results are sharp. In particular, there is an $\eta>0$ such that for $\xi\in B_{(X,d)}\left(\zeta,\left(\frac{\eta}{r_1},\frac{\eta}{r_2}\right)\right)$, we have

$$K_{r_1,r_2}\left(\zeta,\xi\right)\approx \mathrm{Vol}\left(B_{(X,d)}\left(\zeta,\left(\frac{1}{r_1},\frac{1}{r_2}\right)\right)\right)^{-1}.$$

Using these results, one can study maximal operators. Indeed, define three maximal operators:

$$\begin{aligned}
\mathcal{M}f\left(\zeta\right) &= \sup_{\delta_1,\delta_2>0}\frac{1}{\mathrm{Vol}\left(B_{(X,d)}\left(\zeta,\left(\delta_1,\delta_2\right)\right)\right)}\int_{B_{(X,d)}\left(\zeta,\left(\delta_1,\delta_2\right)\right)}\left|f\left(\xi\right)\right|\,d\xi,\\
\mathcal{M}_Lf\left(\zeta\right) &= \sup_{\delta>0}\frac{1}{\mathrm{Vol}\left(B_{(X,d)}\left(\zeta,\left(\delta,0\right)\right)\right)}\int_{B_{(X,d)}\left(\zeta,\left(\delta,0\right)\right)}\left|f\left(\xi\right)\right|\,d\xi,\\
\mathcal{M}_Rf\left(\zeta\right) &= \sup_{\delta>0}\frac{1}{\mathrm{Vol}\left(B_{(X,d)}\left(\zeta,\left(0,\delta\right)\right)\right)}\int_{B_{(X,d)}\left(\zeta,\left(0,\delta\right)\right)}\left|f\left(\xi\right)\right|\,d\xi.
\end{aligned}$$

The results above show

$$\mathcal{M}f\lesssim \mathcal{M}_L\mathcal{M}_Rf.$$

However, it is well-known that the one-parameter maximal functions $\mathcal{M}_L$ and $\mathcal{M}_R$ are bounded on L^p $(1<p\leq\infty)$; due to the fact that $B_{(X,d)}\left(\zeta,\left(\cdot,0\right)\right)$ and $B_{(X,d)}\left(\zeta,\left(0,\cdot\right)\right)$ give rise to spaces of homogeneous type. It follows, then, that $\mathcal{M}$ is also bounded on L^p $(1<p\leq\infty)$.

The goal of Section 6 is to see how far this proof (in its entirety) can be generalized. It was used heavily in [28] that the left invariant vector fields commuted with the right invariant vector fields. In Section 6 we will see that we do not need the relevant vector fields to commute, but can instead just assume that they "almost commute." This is made precise in Section 6.

It is extremely likely maximal results hold for a larger class of our multi-parameter balls than what is shown in Section 6–but the study of unit operators seems very tied to the (rather strong) assumptions in Section 6. We content ourselves, in this paper, with studying maximal operators under these hypotheses. It would be interesting to generalize these results further.

1.2.5. The classical theorem of Frobenius

In this section, we remind the reader of the statement of the theorem of Frobenius. We keep the exposition brief since we use the classical theorem of Frobenius only tangentially in this paper, and this section is more to fix terminology. Suppose M is a connected manifold, and $X_1, \ldots, X_q$ are C^∞ vector fields on M. Suppose, further, that for each i, j there exist C^∞ functions $c_{i,j}^k$ such that:

$$[X_i, X_j] = \sum_k c_{i,j}^k X_k. \tag{1.7}$$

Conditions like (1.7) are referred to as "integrability conditions". In this case, we have the classical theorem of Frobenius:

Theorem 1.6. *For each $x \in M$, there exists a unique, maximal, connected, injectively immersed submanifold $L \subseteq M$ such that:*

- *$x \in L$,*
- *For each $y \in L$, $T_y L = \operatorname{span}\{X_1(y), \ldots, X_q(y)\}$.*

L is called a "leaf".

Remark 1.7. Often, one sees an additional assumption in Theorem 1.6. Namely, that $\dim \operatorname{span}\{X_1, \ldots, X_q\}$ is constant. This assumption is not necessary, and the usual proofs (for instance, the one in [2]) give the stronger result in Theorem 1.6. This was noted in [8].

Remark 1.8. Let $\mathcal{D}$ be a C^∞ module of vector fields on an open set $\Omega \subseteq \mathbb{R}^n$. We call $\mathcal{D}$ a (generalized) distribution. Suppose that $\mathcal{D}$ satisfies two conditions:

1. $\mathcal{D}$ is involutive. That is, if $X, Y \in \mathcal{D}$, then $[X, Y] \in \mathcal{D}$.
2. $\mathcal{D}$ is locally finitely generated as a C^∞ module. That is, for each $x \in \Omega$, there is a neighborhood U containing x such that there exist a finite set of vector fields $X_1, \ldots, X_q \in \mathcal{D}$ such that every $Y \in \mathcal{D}$, when restricted to U, can be written as a linear combination (with coefficients in C^∞) of $X_1, \ldots, X_q$ on U.

Note, under the above hypotheses, $X_1, \ldots, X_q$ satisfy (1.7) on U (since $[X_i, X_j] \in \mathcal{D}$). Thus, one may apply Theorem 1.6 to foliate Ω into leaves, with each leaf L satisfying $T_y L = \mathcal{D}_y$, $\forall y \in L$. Often, the Frobenius theorem is stated in terms of such an "involutive disribution which is locally finitely generated as a C^∞ module," instead of stated in terms of an explicit choice of generators as we have done in Theorem 1.6. The reason we have chosen to state the theorem with an explicit choice of generators is that we will need to discuss how various constants depend on the generators.

Remark 1.9. Above we have stated the result assuming the vector fields are C^∞. In fact, an analogous result (using, again, the usual proofs) holds only assuming that the vector fields are C^1. In fact, there are even results when the vector fields are assumed to be merely Lipschitz (see [23]). However C^1 will be sufficient for our purposes (and most of the applications we have in mind require only C^∞).

We close this section with a discussion of the relationship between Theorem 1.6 and Theorem 1.3. As we mentioned before, Theorem 1.3 implies Theorem 1.6. To understand the philosophy behind Theorem 1.3, let $\mathcal{I}$ be an index set, and suppose for each $\alpha \in \mathcal{I}$ we are given C^∞ vector fields $X_1^\alpha, \ldots, X_q^\alpha$ on a fixed open set Ω. Here, both q and Ω are independent of α. Suppose further that for every $\alpha \in \mathcal{I}$ we have,

$$\left[X_i^\alpha, X_j^\alpha\right] = \sum_k c_{i,j}^{k,\alpha} X_k^\alpha.$$

Suppose, finally, that as α varies over $\mathcal{I}$, X_j^α and $c_{i,j}^{k,\alpha}$ vary over bounded (and therefore pre-compact) subsets of C^∞. Since Theorem 1.6 applies for each $\alpha \in \mathcal{I}$, one might hope that it applies uniformly[17] for $\alpha \in \mathcal{I}$. Indeed, this is the case, and is essentially the statement of Theorem 1.3. Hence, Theorem 1.3 may be informally restated as saying that the theorem of Frobenius holds "uniformly on compact sets" in the above sense.

As it turns out, the classical proofs of Theorem 1.6 do not work uniformly in α in the above sense (this is discussed in Remark 3.4). If we fix $x_0 \in \Omega$ and define $n_0^\alpha = \dim \operatorname{span}\left\{X_1^\alpha(x_0), \ldots, X_q^\alpha(x_0)\right\}$, then the classical proofs also depend on a lower bound for

$$\left| \det_{n_0^\alpha \times n_0^\alpha} X^\alpha(x_0) \right|,$$

which may not be bounded below uniformly for $\alpha \in \mathcal{I}$.[18]

[17] We mean uniformly in the sense that the coordinate charts which define the leaves can be chosen to satisfy good estimates which are uniform in α.

[18] It is not a coincidence that the failure of the classical proofs of the theorem of Frobe-

There is another way to view Theorem 1.3 in relation to Theorem 1.6. Let $X_1, \ldots, X_q$ be C^∞ vector fields satisfying (1.7). Notice we have not assumed that $n_0(x) = \dim \mathrm{span}\{X_1(x), \ldots, X_q(x)\}$ is constant in x. The foliation associated to the involutive distribution generated by $X_1, \ldots, X_q$ is called "singular" if $n_0(x)$ is not constant in x; and if $n_0(x)$ is not constant near a point x_0 then x_0 is called a singular point.

In the classic proofs of Theorem 1.6, the coordinate charts defining the leaves degenerate as one approaches a singular point. Theorem 1.3 avoids this. This is an essential point in Section 6.2.

2. Basic definitions

Fix, for the rest of the paper, a connected open set $\Omega \subseteq \mathbb{R}^n$. Suppose we are given a list of C^1 vector fields $X_1, \ldots, X_q$ defined on Ω, and let X denote this list. As mentioned in Section 1.1, we will often identify this list with the $n \times q$ matrix whose columns are given by the vector fields $X_1, \ldots, X_q$. In addition, we will define (when it makes sense) X^α, where α is an ordered multi-index, in the usual way.[19] Thus, X^α is an $|\alpha|$th order partial differential operator. In the introduction, we defined the Carnot-Carathéodory ball of unit radius centered at $x_0 \in \Omega$. We denoted this ball by $B_X(x_0)$.

It will often be convenient to assume that $B_X(x_0)$ lies "inside" of Ω. More precisely, we make the following definition:

Definition 2.1. Given $x_0 \in \Omega$, we say X satisfies $\mathcal{C}(x_0)$ if for every $a = (a_1, \ldots, a_q) \in (L^\infty([0,1]))^q$, with:

$$\||a|\|_{L^\infty([0,1])} = \left\| \Big(\sum_{j=1}^q |a_j|^2 \Big)^{\frac{1}{2}} \right\|_{L^\infty([0,1])} < 1,$$

there exists a solution $\gamma : [0,1] \to \Omega$ to the ODE:

$$\gamma'(t) = \sum_{j=1}^q a_j(t) X_j(\gamma(t)), \quad \gamma(0) = x_0.$$

Note, by Gronwall's inequality, when this solution exists, it is unique.

nius to be uniform in an appropriate sense lies in the use of a lower bound of a determinant, just as in the work of Nagel, Stein, and Wainger (see Section 1.2.1). Indeed, these two issues are closely related.

[19]For instance, if α were the list $(1,2,1,3)$, then $X^\alpha = X_1X_2X_1X_3$ and $|\alpha| = 4$, the length of the list.

As in the introduction, to define Carnot-Carathéodory balls of (possibly multi-parameter) radii, we assign to each vector field X_j a formal degree $0 \neq d_j \in [0,\infty)^\nu$. Here $\nu \in \mathbb{N}$ is a fixed number, independent of j, representing the number of parameters. We denote the list $(X_1, d_1), \ldots, (X_q, d_q)$ by (X, d). In the introduction, we defined (for $\delta \in [0,\infty)^\nu$) the list $\delta^d X$ to be the list of vector fields $\delta^{d_1} X_1, \ldots, \delta^{d_q} X_q$. Then, we defined the multi-parameter Carnot-Carathéodory ball $B_{(X,d)}(x_0, \delta) := B_{\delta X}(x_0)$. Just as in Definition 2.1 it will often be useful to assume $B_{(X,d)}(x_0, \delta)$ lies "inside" of Ω, and so we make the following definition:

Definition 2.2. Given $x_0 \in \Omega$ and $\delta \in [0,\infty)^\nu$, we say (X, d) satisfies $\mathcal{C}(x_0, \delta)$ if $\delta^d X$ satisfies $\mathcal{C}(x_0)$.

In addition to the balls $B_{(X,d)}(x_0, \delta)$ it will be useful to define some smaller balls. Given $x_0 \in \Omega$ and $\delta \in [0,\infty)^\nu$, we define

$$\widetilde{B}_{(X,d)}(x_0, \delta) = \left\{ y \in \Omega : \exists a \in \mathbb{R}^q, |a| \leq 1, y = \exp\left(a \cdot \delta^d X\right) x_0 \right\}.$$

Note that $\widetilde{B}_{(X,d)}(x_0, \delta) \subseteq B_{(X,d)}(x_0, \delta)$.

Given a list of vector fields along with formal degrees (X, d) and $J = (j_1, \ldots, j_{n_0}) \in \mathcal{I}(n_0, q)$, we defined in the introduction the list of vector fields with formal degrees $(X, d)_J$ and the list of vector fields X_J. Namely, $(X, d)_J$ is the list $(X_{j_1}, d_{j_1}), \ldots, \left(X_{j_{n_0}}, d_{j_{n_0}}\right)$ and X_J is the list $X_{j_1}, \ldots, X_{j_{n_0}}$, while d_J is the list $d_{j_1}, \ldots, d_{j_{n_0}}$.

Note that if (X, d) satisfies $\mathcal{C}(x_0, \delta)$, then so does $(X, d)_J$. In addition, we have,

$$B_{(X,d)_J}(x_0, \delta) \subseteq B_{(X,d)}(x_0, \delta), \quad \widetilde{B}_{(X,d)_J}(x_0, \delta) \subseteq \widetilde{B}_{(X,d)}(x_0, \delta).$$

Often, it will be convenient for our estimates to state the definition of $B_{(X,d)}(x_0, \delta)$ in a slightly different way. Thus, given the formal degrees $d_1, \ldots, d_q$ and given a $a = (a_1, \ldots, a_q) \in \mathbb{R}^q$, $\delta \in [0,\infty)^\nu$, we define:

$$\delta^d a = \left(\delta^{d_1} a_1, \ldots, \delta^{d_q} a_q\right),$$
$$\delta^{-d} a = \left(\delta^{-d_1} a_1, \ldots, \delta^{-d_q} a_q\right).$$

Then we have:

$$B_{(X,d)}(x_0, \delta) = \big\{ y \in \Omega : \exists \gamma : [0,1] \to \Omega, \gamma(0) = x_0, \gamma(1) = y$$
$$\gamma'(t) = a(t) \cdot X(\gamma(t)), \left\| \left| \delta^{-d} a \right| \right\|_{L^\infty([0,1])} < 1 \big\}.$$

3. The (uniform) theorem of Frobenius

In this section, we present a uniform version of the theorem of Frobenius: the special case of Theorem 1.3 when the vector fields are assumed to be linearly independent. The work in this section was heavily influenced by the methods in Section 4 of [30] and those in [22]. In fact, a result similar to a special case of Theorem 3.1 is contained in [30], though the result there is stated somewhat differently (see Section 5.2.1 for a discussion of their results). Our goal, in this section, is to rephrase and generalize the proof methods from these two papers to suit our needs.

In our context, we are faced with a few difficulties not addressed in [30]. A main difficulty we face is that we will not assume an *a priori* smoothness that was assumed in that paper. This will require us to provide a more detailed study of an ODE that arises in that paper. This difference in difficulty here, is that while in that paper existence for a certain ODE was proved via the contraction mapping principle, we must also prove smooth dependence on parameters. Furthermore, we will generalize their results to vector fields that do not necessarily span the tangent space. While this may seem like an artificial generalization, it will prove to be essential to our study of maximal functions and unit operators in Sections 4.2 and 6. Finally, we must also combine these methods with the methods in [22] to prove the relationships between the various balls we will define.

Let $X = (X_1, \ldots, X_{n_0})$ be n_0 C^1 vector fields with single-parameter formal degrees $d = (d_1, \ldots, d_{n_0}) \in (0, \infty)^{n_0}$ defined on the fixed connected open set $\Omega \subseteq \mathbb{R}^n$. Fix $1 \geq \xi > 0$, $x_0 \in \Omega$, and suppose that (X, d) satisfies $\mathcal{C}(x_0, \xi)$. Suppose further that the X_js satisfy an integrability condition on $B_{(X,d)}(x_0, \xi)$ given by:

$$[X_j, X_k] = \sum_l c_{j,k}^l X_l.$$

In this section, we will assume that:

- $X_1(x_0), \ldots, X_{n_0}(x_0)$ are linearly independent.
- $\|X_j\|_{C^1\left(B_{(X,d)}(x_0,\xi)\right)} < \infty$, for every $1 \leq j \leq n_0$.
- For $|\alpha| \leq 2$, $X^\alpha c_{j,k}^l \in C^0\left(B_{(X,d)}(x_0, \xi)\right)$, and

$$\sum_{|\alpha| \leq 2} \left\| X^\alpha c_{j,k}^l \right\|_{C^0\left(B_{(X,d)}(x_0,\xi)\right)} < \infty, \qquad \text{for all } j, k, l.$$

We will say that C is an admissible constant if C can be chosen to depend only on a fixed upper bound, $d_{max} < \infty$, for $d_1, \ldots, d_{n_0}$, a fixed lower bound

$d_{min} > 0$ for $d_1, \ldots, d_{n_0}$, a fixed upper bound for n (and therefore for n_0), a fixed lower bound, $\xi_0 > 0$, for ξ, and a fixed upper bound for the quantities:

$$\|X_j\|_{C^1\left(B_{(X,d)}(x_0,\xi)\right)}, \quad \sum_{|\alpha|\leq 2} \left\|X^\alpha c_{j,k}^l\right\|_{C^0\left(B_{(X,d)}(x_0,\xi)\right)}.$$

Furthermore, if we say that C is an m-admissible constant, we mean that in addition to the above, we assume that:

$$\sum_{|\alpha|\leq m} \left\|X^\alpha c_{j,k}^l\right\|_{C^0\left(B_{(X,d)}(x_0,\xi)\right)} < \infty,$$

for every j, k, l (in particular, these derivatives up to order m exist and are continuous). C is allowed to depend on m, all the quantities an admissible constant is allowed to depend on, and a fixed upper bound for the above quantity. Note that $\lesssim_0, \lesssim_1, \lesssim_2$, and $\lesssim$ all denote the same thing.

For $\eta > 0$, a sufficiently small admissible constant, define the map:

$$\Phi : B_{n_0}(\eta) \to \widetilde{B}_{(X,d)}(x_0, \xi)$$

by

$$\Phi(u) = \exp(u \cdot X) x_0.$$

Note that, by Theorem A.1, Φ is C^1. The main theorem of this section is the following:

Theorem 3.1. *There exist admissible constants $\eta_1 > 0$, $\xi_1 > 0$, such that:*

- $\Phi : B_{n_0}(\eta_1) \to \widetilde{B}_{(X,d)}(x_0, \xi)$ *is one-to-one.*
- *For all $u \in B_{n_0}(\eta_1)$, $|\det_{n_0\times n_0} d\Phi(u)| \approx |\det_{n_0\times n_0} X(x_0)|$.*
- $B_{(X,d)}(x_0, \xi_1) \subseteq \Phi(B_{n_0}(\eta_1)) \subseteq \widetilde{B}_{(X,d)}(x_0, \xi) \subseteq B_{(X,d)}(x_0, \xi)$.

Furthermore, if we let Y_j be the pullback of X_j under the map Φ, then we have:

$$\|Y_j\|_{C^m\left(B_{n_0}(\eta_1)\right)} \lesssim_m 1 \tag{3.1}$$

in particular,

$$\|Y_j\|_{C^2\left(B_{n_0}(\eta_1)\right)} \lesssim 1.$$

Finally, if for $u \in B_{n_0}(\eta_1)$ we define the $n_0 \times n_0$ matrix $A(u)$ by:[20]

$$(Y_1, \ldots, Y_{n_0}) = (I + A) \bigtriangledown_u$$

then,

$$\sup_{u \in B_{n_0}(\eta_1)} \|A(u)\| \leq \frac{1}{2}.$$

[20]Here we are thinking of $\bigtriangledown_u$ as the vector $(\partial_{u_1}, \ldots, \partial_{u_{n_0}})$.

This section will be devoted to the proof of Theorem 3.1.

Remark 3.2. In [30], the map Φ was defined with a large parameter K. Then, a result like (3.1) was proven by taking K large depending on m. It is important for the applications we have in mind that this procedure is not necessary. In our setup, this procedure is similar to taking the parameter κ in Theorem 3.10 small depending on m; however we will see that we will be able to fix $\kappa = \frac{1}{2}$ throughout.

Remark 3.3. The formal degrees, $d_1, \ldots, d_{n_0}$ do not play an essential role in this section. Indeed note that they do not play a role in the assumptions for Theorem 3.1. Moreover, since $\xi_1, \xi \approx 1$, they do not play a role in the conclusion either. Indeed, Theorem 3.1 with any choice of $d_1, \ldots, d_{n_0} \in (0, \infty)$ is equivalent to the theorem with any other choice (though the various constants in the conclusion of Theorem 3.1 will depend on the choice of the ds). The reason we have chosen to state Theorem 3.1 with an arbitrary choice of ds (instead of taking, say, $d_1 = \cdots = d_{n_0} = 1$) is that when we prove Theorem 5.3 we will be, in effect, applying Theorem 3.1 infinitely many times. Having stated Theorem 3.1 for general d will allow us to seamlessly apply the results here without any hand-waving about how various constants depend on the formal degrees.

Remark 3.4. As was discussed in Section 1.2.1, the methods in [22] fail to prove Theorem 3.1. It is also worth noting that the methods usually used to prove the theorem of Frobenius are insufficient to prove Theorem 3.1. For simplicity, we discuss the proof in [14], but similar remarks hold for all previous proofs we know of. In [14], an invertible linear transformation was applied to $X_1, \ldots, X_{n_0}$ (call the resulting vector fields $V_1, \ldots, V_{n_0}$). This was done in such a way that $[V_i, V_j] = 0$ for every i, j. Because of this, the map:

$$u \mapsto e^{u \cdot V} x_0$$

is easy to study. Unfortunately, we know of no *a priori* way to create such an invertible linear transformation without destroying the admissible constants. *A fortiori*, however, we may just push forward the linear transformation $(I + A)^{-1}$ via the map Φ to obtain such a linear transformation. This idea seems to yield no nontrivial new information.

Remark 3.5. Morally, Theorem 3.1 (along with Theorems 4.1 and 5.3) is a compactness result.

This is discussed at the end of Section 1.2.5. The use of this compactness can be seen every time we apply Theorem A.3. Moreover, this compactness perspective was taken up in Section 4 of [28]. In fact, one of the main consequences of this paper is that one may remove condition 4 of Definition 4.4

of [28], and still obtain the relevant results (this is tantamount to saying that we do not require a lower bound for a determinant as discussed in Section 1.2.1). Thus, from the remarks in that paper, one can easily see the results in this paper from the perspective of compactness.

The next two lemmas we state in slightly greater generality than we need, since we will refer to the proofs later in the paper.

Lemma 3.6. *Fix* $1 \leq n_1 \leq n_0$. *Then, for* $1 \leq j \leq n_0$, $I \in \mathcal{I}(n_1, n)$, $J \in \mathcal{I}(n_1, n_0)$, $x \in B_{(X,d)}(x_0, \xi)$,

$$\left|X_j \det X(x)_{I,J}\right| \lesssim \left|\det_{n_1 \times n_1} X(x)\right|.$$

Proof. We use the notation $\mathcal{L}_U$ to denote the Lie derivative with respect to the vector field U, and i_V to denote the interior product with the vector field V. $\mathcal{L}_U$ and i_V have the following, well-known, properties:

- $\mathcal{L}_U f = Uf$ for functions f.
- $[\mathcal{L}_U, i_V] = i_{[U,V]}$.
- $\mathcal{L}_U \omega = i_U d\omega + d i_U \omega$, for forms ω.
- $\mathcal{L}_U(\omega_1 \wedge \omega_2) = (\mathcal{L}_U \omega_1) \wedge \omega_2 + \omega_1 \wedge (\mathcal{L}_U \omega_2)$ for forms ω_1, ω_2.
- If $U = \sum_k b_k \frac{\partial}{\partial x_k}$, then,

$$\mathcal{L}_U dx_k = d i_U dx_k = db_k = \sum \frac{\partial b_k}{\partial x_j} dx_j.$$

Fix $I = (i_1, \ldots, i_{n_1}), J = (j_1, \ldots, j_{n_1})$ as in the statement of the lemma. Then,

$$\det X(x)_{I,J} = i_{X_{j_{n_1}}} i_{X_{j_{n_1-1}}} \cdots i_{X_{j_1}} dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_{n_1}}.$$

Thus, we see:

$$\begin{aligned}
X_j \det X(x)_{I,J} &= \mathcal{L}_{X_j} i_{X_{j_{n_1}}} i_{X_{j_{n_1-1}}} \cdots i_{X_{j_1}} dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_{n_1}} \\
&= i_{\left[X_j, X_{j_{n_1}}\right]} i_{X_{j_{n_1-1}}} \cdots i_{X_{j_1}} dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_{n_1}} \\
&\quad + i_{X_{j_{n_1}}} i_{\left[X_j, X_{j_{n_1-1}}\right]} \cdots i_{X_{j_1}} dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_{n_1}} \\
&\quad + \cdots + i_{X_{j_{n_1}}} i_{X_{j_{n_1-1}}} \cdots i_{\left[X_j, X_{j_1}\right]} dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_{n_1}} \\
&\quad + i_{X_{j_{n_1}}} i_{X_{j_{n_1-1}}} \cdots i_{X_{j_1}} \mathcal{L}_{X_j} \left(dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_{n_1}}\right).
\end{aligned} \tag{3.2}$$

Every term, except the last term, on the RHS of (3.2) is easy to estimate. We do the first term as an example, and all of the others work in the same way:

$$\begin{aligned}&\left|i_{[X_j,X_{j_{n_1}}]}i_{X_{j_{n_1-1}}}\cdots i_{X_{j_1}}dx_{i_1}\wedge dx_{i_2}\wedge\cdots\wedge dx_{i_{n_1}}\right| =\\ &\quad=\left|\sum_{k=1}^{n_0}c^k_{j,j_{n_1}}i_{X_k}i_{X_{j_{n_1-1}}}\cdots i_{X_{j_1}}dx_{i_1}\wedge dx_{i_2}\wedge\cdots\wedge dx_{i_{n_1}}\right|\lesssim\left|\det_{n_1\times n_1}X(x)\right|.\end{aligned}$$

Since, for each k, $i_{X_k}i_{X_{j_{n_1-1}}}\cdots i_{X_{j_1}}dx_{i_1}\wedge dx_{i_2}\wedge\cdots\wedge dx_{i_{n_1}}$ is either 0 or of the form $\pm\det X(x)_{I,J'}$ for some $J'\in\mathcal{I}(n_1,n_0)$.

We now turn to the last term on the RHS of (3.2). We have:

$$\begin{aligned}&\mathcal{L}_{X_j}\left(dx_{i_1}\wedge dx_{i_2}\wedge\cdots\wedge dx_{i_{n_1}}\right)=\\ &\quad=\left(\mathcal{L}_{X_j}dx_{i_1}\right)\wedge dx_{i_2}\wedge\cdots\wedge dx_{i_{n_1}}+dx_{i_1}\wedge\left(\mathcal{L}_{X_j}dx_{i_2}\right)\wedge\cdots\wedge dx_{i_{n_1}}\\ &\qquad+\cdots+dx_{i_1}\wedge dx_{i_2}\wedge\cdots\wedge\left(\mathcal{L}_{X_j}dx_{i_{n_1}}\right).\end{aligned}$$

So we may separate the last term on the RHS of (3.2) into a sum of n_1 terms. We bound just the first, the bounds of the others being similar. To to this, let $X_j=\sum_k b^k_j\frac{\partial}{\partial x_k}$. Note that $\left\|b^k_j\right\|_{C^1\left(B_{(X,d)}(x_0,\xi)\right)}\lesssim 1$.

$$\begin{aligned}&\left|i_{X_{j_{n_1}}}i_{X_{j_{n_1-1}}}\cdots i_{X_{j_1}}\left(\mathcal{L}_{X_j}dx_{i_1}\right)\wedge dx_{i_2}\wedge\cdots\wedge dx_{i_{n_1}}\right|=\\ &\quad=\left|\sum_l\frac{\partial b^{i_1}_j}{\partial x_l}i_{X_{j_{n_1}}}i_{X_{j_{n_1-1}}}\cdots i_{X_{j_1}}dx_l\wedge dx_{i_2}\wedge\cdots\wedge dx_{i_{n_1}}\right|\\ &\quad\lesssim\left|\det_{n_1\times n_1}X(x)\right|.\end{aligned}$$

since each of the terms $i_{X_{j_{n_1}}}i_{X_{j_{n_1-1}}}\cdots i_{X_{j_1}}dx_l\wedge dx_{i_2}\wedge\cdots\wedge dx_{i_{n_1}}$ is either 0 or of the form $\pm\det X(x)_{I',J}$ for some $I'\in\mathcal{I}(n_1,n)$. ■

Remark 3.7. The reader wishing to avoid the use of Lie derivatives in Lemma 3.6 should consult Lemma 2.6 of [22] where a similar result in the special case $n_1=n$ is shown directly, without the use of Lie derivatives. However, the proof we give in Lemma 3.6 is easily adapted to other situations that will arise in this paper (e.g. Lemmas 4.10 and 4.13), while the proof in [22] becomes progressively more complicated to generalize.

Lemma 3.8. *For $y\in B_{(X,d)}(x_0,\xi)$, $1\leq n_1\leq n_0$, we have*

$$\left|\det_{n_1\times n_1}X(y)\right|\approx\left|\det_{n_1\times n_1}X(x_0)\right|$$

In particular, since $|\det_{n_0\times n_0}X(x_0)|\neq 0$, $|\det_{n_0\times n_0}X(y)|\neq 0$.

Proof. Since $y \in B_{(X,d)}(x_0, \xi)$, there exists $\gamma : [0,1] \to B_{(X,d)}(x_0, \xi)$ with

- $\gamma(0) = x_0$, $\gamma(1) = y$,
- $\gamma'(t) = a(t) \cdot X(\gamma(t))$,
- $a \in (L^\infty([0,1]))^{n_0}$,
- $\left\| \left| \xi^{-d} a \right| \right\|_{L^\infty([0,1])} < 1$.

But, then consider:

$$\begin{aligned} \frac{d}{dt} \left| \det_{n_1 \times n_1} X(\gamma(t)) \right|^2 &= 2 \sum_{\substack{I \in \mathcal{I}(n_1, n) \\ J \in \mathcal{I}(n_1, n_0)}} \det X_{I,J}(\gamma(t)) \frac{d}{dt} \det X_{I,J}(\gamma(t)) \\ &= 2 \sum_{\substack{I \in \mathcal{I}(n_1, n) \\ J \in \mathcal{I}(n_1, n_0)}} \det X_{I,J}(\gamma(t)) ((a \cdot X) \det X_{I,J})(\gamma(t)) \\ &\lesssim \left| \det_{n_1 \times n_1} X(\gamma(t)) \right|^2 \end{aligned}$$

where, in the last step, we have applied Lemma 3.6. Hence, Gronwall's inequality shows:

$$\left| \det_{n_1 \times n_1} X(y) \right| = \left| \det_{n_1 \times n_1} X(\gamma(1)) \right| \lesssim \left| \det_{n_1 \times n_1} X(\gamma(0)) \right| = \left| \det_{n_1 \times n_1} X(x_0) \right|.$$

Reversing the path γ and applying the same argument, we see that:

$$\left| \det_{n_1 \times n_1} X(x_0) \right| \lesssim \left| \det_{n_1 \times n_1} X(y) \right|,$$

completing the proof. ∎

Now consider the map $\Phi : B_{n_0}(\eta) \to B_{(X,d)}(x_0, \xi)$. $d\Phi(0) = X(x_0)$, and it follows that $\det_{n_0 \times n_0} d\Phi(0) \neq 0$. Hence, if we consider Φ as a map to the leaf generated by X passing through the point x_0, the inverse function theorem shows that there is a (non-admissible) $\delta > 0$ such that:

$$\Phi : B_{n_0}(\delta) \to \Phi(B_{n_0}(\delta))$$

is a C^1 diffeomorphism. Pullback the vector field X_j via the map Φ to $B_{n_0}(\delta)$. Call this C^0 vector field $\widehat{Y}_j$.

Clearly $\widehat{Y}_j(0) = \frac{\partial}{\partial u_j}$. Write:

$$\widehat{Y}_j = \frac{\partial}{\partial u_j} + \sum_k \hat{a}_j^k \frac{\partial}{\partial u_k} \tag{3.3}$$

with $\hat{a}_j^k(0) = 0$. Moreover, in polar coordinates, for ω fixed, Remark A.2 shows that $\hat{a}_j^k(r\omega)$ is C^1 in the r variable, and it follows that for ω fixed, $\hat{a}_j^k(r\omega) = O(r)$. We will now show that $\hat{a}_j^k$ satisfies an ODE in the r variable. The derivation of this ODE is classical (see, for instance, page 155 of [2], though we follow the presentation of [30]), and is the main starting point for this entire section. We include the derivation here, since it is not very long, and is of fundamental importance to the rest of the paper.

Continuing in polar coordinates,

$$\Phi(r,\omega) = \exp(r(\omega\cdot X))\, x_0.$$

Hence,

$$d\Phi(r\partial_r)(\Phi(r,\omega)) = r d\Phi(\partial_r)(\Phi(r,\omega)) = r\omega\cdot X(\Phi(r,\omega)).$$

Writing this in Cartesian coordinates, we have the following vector field identity on $B_{n_0}(\delta)$:

$$\sum_{j=1}^{n_0} u_j \frac{\partial}{\partial u_j} = \sum_{j=1}^{n_0} u_j \widehat{Y}_j. \tag{3.4}$$

Taking the lie bracket of (3.4) with $\widehat{Y}_i$, we obtain:

$$\begin{aligned}\sum_{j=1}^{n_0}\left(\widehat{Y}_i(u_j)\,\partial_{u_j} + u_j\left[\widehat{Y}_i,\partial_{u_j}\right]\right) &= \sum_{j=1}^{n_0}\left(\widehat{Y}_i(u_j)\,\widehat{Y}_j + u_j\left[\widehat{Y}_i,\widehat{Y}_j\right]\right)\\ &= \sum_{j=1}^{n_0}\left(\widehat{Y}_i(u_j)\,\widehat{Y}_j + u_j\sum_{l=1}^{n_0}\tilde{c}_{i,j}^{l}(u)\,\widehat{Y}_l\right),\end{aligned} \tag{3.5}$$

where $\tilde{c}_{i,j}^k(u) = c_{i,j}^k(\Phi(u))$, and we have used the fact that

$$\left[\widehat{Y}_i,\widehat{Y}_j\right] = \sum \tilde{c}_{i,j}^k \widehat{Y}_k.$$

Remark 3.9. Since $\widehat{Y}_i$ is not C^1, one might worry about our manipulations in (3.5). This turns out to not be a problem. Indeed, it makes sense to take the above commutator, since $\widehat{Y}_i$ is C^1 in the r variable (and we are commuting it with $r\partial_r$). Then, the computations on the LHS of (3.5) may be done in the sense of distributions, while the computations on the RHS may be done by pushing everything forward via the map Φ. We leave the details to the reader.

We re-write (3.5) as:

$$\left(\sum_{j=1}^{n_0} u_j \left[\partial_{u_j}, \widehat{Y}_i - \partial_{u_i}\right]\right) + \widehat{Y}_i - \partial_{u_i} = \tag{3.6}$$
$$= -\left(\sum_{j=1}^{n_0} \left(\widehat{Y}_i - \partial_{u_i}\right)(u_j)\left(\widehat{Y}_j - \partial_{u_j}\right)\right) - \sum_{j=1}^{n_0}\sum_{l=1}^{n_0} u_j \tilde{c}_{i,j}^{l}(u)\,\widehat{Y}_l.$$

Plugging (3.3) into (3.6) we have:

$$\sum_{j=1}^{n_0}\sum_{k=1}^{n_0} u_j\left(\partial_{u_j}\hat{a}_j^k\right)\partial_{u_k} + \sum_{k=1}^{n_0}\hat{a}_i^k\partial_{u_k} = \tag{3.7}$$
$$= -\left(\sum_{j=1}^{n_0}\sum_{k=1}^{n_0}\hat{a}_i^j\hat{a}_j^k\partial_{u_k}\right) - \sum_{k=1}^{n_0}\left(\sum_{j=1}^{n_0} u_j\tilde{c}_{i,j}^k\right)\partial_{u_k} - \sum_{l=1}^{n_0}\sum_{k=1}^{n_0}\left(\sum_{j=1}^{n_0} t_j\tilde{c}_{i,j}^l\right)\hat{a}_l^k\partial_{u_k}.$$

Taking the ∂_{u_k} component, and writing $\sum_{j=1}^{n_0} u_j\partial_{u_j} + 1 = \partial_r r$, we have from (3.7):

$$\partial_r r\hat{a}_i^k = -\sum_{j=1}^{n_0}\hat{a}_i^j\hat{a}_j^k - \sum_{j=1}^{n_0} u_j\tilde{c}_{i,j}^k - \sum_{l=1}^{n_0}\left(\sum_{j=1}^{n_0} u_j\tilde{c}_{i,j}^l\right)\hat{a}_l^k. \tag{3.8}$$

Define two $n_0 \times n_0$ matrices, $\widehat{A}, C_u$ by:

$$\widehat{A}_{i,k} := \left(\hat{a}_i^k\right), \quad (C_u)_{i,k} := \left(\sum_{j=1}^{n_0} u_j\tilde{c}_{i,j}^k\right), \quad 1 \leq i,k \leq n_0.$$

Using this, (3.8) may be re-written as the matrix valued ODE:

$$\partial_r r\widehat{A} = -\widehat{A}^2 - C_u\widehat{A} - C_u. \tag{3.9}$$

Theorem 3.10. *Fix $\frac{1}{2} \geq \kappa > 0$ (throughout the paper we will choose $\kappa = \frac{1}{2}$). Consider the differential equation:*

$$\partial_r rA(r\omega) = -A(r\omega)^2 - C_u(r\omega)A(r\omega) - C_u(r\omega), \tag{3.10}$$

defined for $A : B_{n_0}(\eta) \to \mathbb{M}_{n_0\times n_0}(\mathbb{R})$, where $\mathbb{M}_{n_0\times n_0}(\mathbb{R})$ denotes the set of $n_0\times n_0$ real matrices. Then, there exists an admissible constant $\eta_1 = \eta_1(\kappa) > 0$ such that there exists a unique solution $A \in C(B_{n_0}(\eta_1); \mathbb{M}_{n_0\times n_0}(\mathbb{R}))$ to (3.10) satisfying $A(r\omega) = O(r)$ for each fixed ω. Moreover, this solution satisfies:

- $\|A(t)\| \lesssim |t|$.
- $\sup_{t\in B_{n_0}(\eta_1)} \|A(t)\| \leq \kappa$.

Furthermore, if $\tilde{c}_{i,j}^k \in C^m(B_{n_0}(\eta_1))$ *with* $\left\|\tilde{c}_{i,j}^k\right\|_{C^m\left(B_{n_0}(\eta_1)\right)} < \infty$*, then* $A \in C^m(B_{n_0}(\eta); \mathbb{M}_{n_0\times n_0}(\mathbb{R}))$*, and if* $\widetilde{C}_{m,\eta_1}$ *is a fixed upper bound for:*

$$\left\|\tilde{c}_{i,j}^k\right\|_{C^m\left(B_{n_0}(\eta_1)\right)}, \quad 1 \leq i, j, k \leq n_0,$$

then, there exists an admissible constant $C_m = C_m\big(m, \widetilde{C}_{m,\eta_1}\big)$ *such that:*

$$\|A\|_{C^m\left(B_{n_0}(\eta_1); \mathbb{M}_{n_0\times n_0}(\mathbb{R})\right)} \leq C_m. \tag{3.11}$$

Note that (3.10) is not a standard ODE (due to the factor of r on the left hand side), and so we cannot apply the standard theorems for existence and dependence on parameters. Fortunately, though, we will be able to prove Theorem 3.10, by adapting the methods of [10]. In [30], the solution A was assumed to be *a priori* C^∞, thereby removing many of the difficulties in the proof of Theorem 3.10. Before we begin the proof, we need two preliminary lemmas:

Lemma 3.11. *Fix* $\epsilon > 0$*. Suppose* $g \in C^m(B_{n_0}(\epsilon))$*. Define* h *on* $B_{n_0}(\epsilon)$ *by:*

$$h(r\omega) = \begin{cases} \frac{1}{r}\int_0^r g(s\omega)\, ds & \text{if } r \neq 0, \\ g(0) & \text{if } r = 0. \end{cases} \tag{3.12}$$

Then, $h \in C^m(B_{n_0}(\epsilon))$*. Moreover, if* α *is a multi-index with* $|\alpha| \leq m$*, we have:*

$$(\partial_u^\alpha h)(r\omega) = \begin{cases} \frac{1}{r^{|\alpha|+1}}\int_0^r s^{|\alpha|}(\partial_u^\alpha g)(s\omega)\, ds & \text{if } r \neq 0, \\ \frac{1}{|\alpha|+1}(\partial_u^\alpha g)(0) & \text{if } r = 0. \end{cases} \tag{3.13}$$

Proof. Note that, since $g \in C^m$, the right hand sides of (3.12) and (3.13) are both continuous in r. Note, also, that to prove the lemma, it suffices to prove the formula (3.13) for $g \in C^\infty$, as then the linear map $g \mapsto h$ will extend as a map $C^\infty \to C^m$ to a map $C^m \to C^m$. Hence, we prove the lemma just under the assumption $g \in C^\infty$ (this reduction is not necessary for our proof, but it simplifies notation a bit).

First, we prove the lemma for $r \neq 0$. Away from $r = 0$, h is clearly C^∞, and so we need only verify the formula (3.13). h satisfies the formula:

$$\partial_r r h(r\omega) = g(r\omega).$$

Apply ∂_u^α to both sides of this formula. Using the fact that $[\partial_u^\alpha, \partial_r r] = |\alpha|\, \partial_u^\alpha$, we have:

$$\partial_r r(\partial_u^\alpha h)(r\omega) + |\alpha|(\partial_u^\alpha h)(r\omega) = (\partial_u^\alpha g)(r\omega).$$

Multiplying both sides by $r^{|\alpha|}$, we obtain:

$$\partial_r r^{|\alpha|+1} \left(\partial_u^\alpha h\right)(r\omega) = r^{|\alpha|} \left(\partial_u^\alpha g\right)(r\omega)$$

and (3.13) follows for $r \neq 0$.

Hence, to complete the proof, we need only show that $\partial_u^\alpha h$ exists at 0 and is given by the $\frac{1}{|\alpha|+1} \left(\partial_u^\alpha g\right)(0)$. We first consider the case when:

$$\partial_u^\beta g(0) = 0, \quad 0 \leq |\beta| \leq m,$$

and we prove the result by induction on the order of α, our base case being the trivial case $|\alpha| = 0$. Thus, suppose we have the result for some α, $|\alpha| < m$ and we wish to show that the following derivative exists, and equals 0:

$$\partial_{u_j} \partial_u^\alpha h(r\omega)\Big|_{r=0} = \partial_{u_j} \begin{cases} \frac{1}{r^{|\alpha|+1}} \int_0^r s^{|\alpha|} \left(\partial_u^\alpha g\right)(s\omega)\, ds & \text{if } r \neq 0 \\ 0 & \text{if } r = 0 \end{cases}\Bigg|_{r=0}$$

And this will follow if we can show that:

$$\frac{1}{r^{|\alpha|+1}} \int_0^r s^{|\alpha|} \left(\partial_u^\alpha g\right)(s\omega)\, ds = o(r). \tag{3.14}$$

But, by our assumption on g, $\left(\partial_u^\alpha g\right)(s\omega) = O\left(s^2\right)$ and (3.14) follows, completing the proof in this case.

Now turn to the general case $g \in C^\infty$. We may write:

$$g(u) = \sum_{|\beta| \leq m} \frac{1}{\beta!} \left(\partial_u^\beta g\right)(0)\, u^\beta + g_e(u)$$

where g_e vanishes to order m at 0. Thus, by linearity of the map $g \mapsto h$, it suffices to prove the lemma for monomials u^β. Since we know (3.13) holds away from $r = 0$ and we know the RHS of (3.13) is continuous, it suffices to show that if $g = u^\beta$, then $h \in C^\infty$. But in this case, $h = \frac{1}{|\beta|+1} u^\beta \in C^\infty$, completing the proof. ∎

Lemma 3.12. ([10, p. 2060]). *Suppose (M, ρ) is a metric space, and suppose $\left(Q_n\right)_{n=0}^\infty$ is a sequence of contractions on M for which there exists a number $c < 1$ such that:*

$$\rho\left(Q_n(x), Q_n(y)\right) \leq c\rho(x, y)$$

for all $x, y \in M$ and all n. Suppose also that there is a point $x_\infty \in M$ such that $Q_n\left(x_\infty\right) \to x_\infty$ as $n \to \infty$. Let $x_0 \in M$ be arbitrary, and define a sequence $\left(x_n\right)$ by setting:

$$x_{n+1} = Q_n\left(x_n\right).$$

Then, $x_n \to x_\infty$ as $n \to \infty$.

Proof of Theorem 3.10. It is easy to see from the definition C_u that:

$$\left\|C_u(r\omega)\right\| \leq Dr$$

where D is an admissible constant. Take $\eta_1 = \eta_1(\kappa) > 0$ to be an admissible constant so small that:

$$\kappa^2 + \frac{D\eta_1}{2}(\kappa+1) \leq \kappa, \quad \kappa + \frac{D\eta_1}{3} \leq \frac{3}{4}.$$

Our first step will be to show the existence of A using the contraction mapping principle. Moreover, this contraction mapping principle may be considered the base case in an induction we will use at the end of the proof, to establish the regularity of A. Consider the metric space:

$$M := \Big\{ A \in C\left(B_{n_0}(\eta_1); \mathbb{M}_{n_0\times n_0}(\mathbb{R})\right) : \ A(0) = 0, \\ \sup_{\substack{0<r\leq\eta_1\\ \omega\in S^{n_0-1}}} \left\|\frac{1}{r}A(r\omega)\right\| < \infty, \sup_{t\in B_{n_0}(\eta_1)} \left\|A(t)\right\| \leq \kappa \Big\}$$

with the metric:

$$\rho(A,B) = \sup_{\substack{0<r\leq\eta_1\\ \omega\in S^{n_0-1}}} \left\|\frac{1}{r}\left(A(r\omega) - B(r\omega)\right)\right\|.$$

Note that M is complete with respect to the metric ρ. Define the map $T: M \to C\left(B_{n_0}(\eta_1); \mathbb{M}_{n_0\times n_0}(\mathbb{R})\right)$, by:

$$TA(r\omega) = \begin{cases} \frac{1}{r}\int_0^r -A(s\omega)^2 - C_u(s\omega)A(s\omega) - C_u(s\omega)\,ds & \text{if } r\neq 0,\\ 0 & \text{if } r = 0.\end{cases}$$

Note that, by Lemma 3.11, $TA \in C\left(B_{n_0}(\eta_1); \mathbb{M}_{n_0\times n_0}(\mathbb{R})\right)$.

Our first goal is to show that $T: M \to M$. Consider, for $0 < r \leq \eta_1$, $\omega \in S^{n_0-1}$, $A \in M$,

$$\begin{aligned}\left\|TA(r\omega)\right\| &\leq \frac{1}{r}\int_0^r \left\|A(s\omega)\right\|^2 + \left\|C_u(s\omega)\right\|\left\|A(s\omega)\right\| + \left\|C_u(s\omega)\right\| ds\\ &\leq \frac{1}{r}\int_0^r \left(\kappa^2 + Ds\kappa + Ds\right) ds\\ &\leq \kappa^2 + \frac{D\eta_1}{2}\kappa + \frac{D\eta_1}{2} \leq \kappa.\end{aligned}$$

Thus, by the definition of TA, $\sup_{t\in B_{n_0}(\eta_1)} \left\|TA(t)\right\| \leq \kappa$.

Next, we have:

$$\left\| \frac{1}{r} TA\left(r\omega\right) \right\| \leq \frac{1}{r^2} \int_0^r \left(s\kappa\rho\left(0, A\right) + Ds\kappa + Ds\right) ds$$
$$= \frac{\kappa}{2}\rho\left(0, A\right) + \frac{D\kappa}{2} + \frac{D}{2} < \infty.$$

Hence, $T : M \to M$.

Next, we wish to show that T is a contraction. Consider, suppressing the dependence on $s\omega$ in the integrals,

$$\left\| \frac{1}{r}\left(TA\left(r\omega\right) - TB\left(r\omega\right)\right) \right\| = \left\| \frac{1}{r^2} \int_0^r -\left(A-B\right)A - B\left(A-B\right) - C_u\left(A-B\right) \right\|$$
$$\leq \frac{1}{r^2} \int_0^r \left(2s\kappa\rho\left(A, B\right) + Ds^2\rho\left(A, B\right)\right) ds$$
$$\leq \kappa\rho\left(A, B\right) + \frac{D\eta_1}{3}\rho\left(A, B\right)$$
$$\leq \frac{3}{4}\rho\left(A, B\right)$$

where the last line follows by our choice of η_1. Thus, we have $\rho\left(TA, TB\right) \leq \frac{3}{4}\rho\left(A, B\right)$.

Applying the contraction mapping principle, there exists a unique fixed point $A \in M$ such that $TA = A$. This is the desired solution to (3.10). Since $A \in M$, we have $\sup_{t \in B_{n_0}(\eta_1)} \|A\left(t\right)\| \leq \kappa$. Moreover, since $A = \lim_{n\to\infty} T^n\left(0\right)$, we have:

$$\rho\left(0, A\right) = \lim_{n\to\infty} \rho\left(0, T^n 0\right) \leq \sum_{n=1}^{\infty} \rho\left(T^{n-1}0, T^n 0\right)$$
$$\leq \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \rho\left(0, T0\right) = 4\rho\left(0, T0\right)$$

and for $r \neq 0$, we have:

$$\left\| \frac{1}{r} T0\left(r\omega\right) \right\| \leq \frac{1}{r^2} \int_0^r Dsds \leq \frac{D}{2}$$

and so $\rho\left(0, T0\right) \lesssim 1$ and therefore $\rho\left(0, A\right) \lesssim 1$. This can be rephrased as $\|A\left(t\right)\| \lesssim |t|$.

We now turn to uniqueness of the solution A. Suppose B is another solution (we are not, necessarily, assuming $B \in M$). Suppose that, for ω fixed, $\|B\left(r\omega\right)\| = O\left(r\right)$. Then, we have:

$$\|r\left(A\left(r\omega\right) - B\left(r\omega\right)\right)\| \leq \int_0^r \left(\|s\left(A - B\right)\| \left[\left\|\frac{A}{s}\right\| + \left\|\frac{B}{s}\right\| + \left\|\frac{C_u}{s}\right\| \right] \right) ds$$

And applying the integral form of Gronwall's inequality to $\|r\left(A - B\right)\|$ shows that $A = B$.

To conclude the proof, we need to show that if $\tilde{c}_{i,j}^k \in C^m$, then $A \in C^m$, and to estimate the C^m norm of A. First, we show that $A \in C^m$. To do this, we will show that $T^n 0 \to A$ in $C^m\left(B_{n_0}\left(\eta_1\right); \mathbb{M}_{n_0\times n_0}\left(\mathbb{R}\right)\right)$ (here we mean the Banach space of those C^m functions all of whose derivatives up to order m are bounded on $B_{n_0}\left(\eta_1\right)$). We proceed by induction on m, our base case being $m=0$, which we have already proven, by the contraction mapping principle. Thus, suppose $\left\|\tilde{c}_{i,j}^k\right\|_{C^m\left(B_{n_0}(\eta_1)\right)} < \infty$ for $1 \leq i,j,k \leq n_0$ and suppose

$$\lim_{n\to\infty} \left\|T^n 0 - A\right\|_{C^{m-1}\left(B_{n_0}(\eta_1); \mathbb{M}_{n_0\times n_0}(\mathbb{R})\right)} = 0.$$

Fix $|\alpha| = m$. We will show that

$$\partial_u^\alpha T^n 0$$

converges in $C^0\left(B_{n_0}\left(\eta_1\right); \mathbb{M}_{n_0\times n_0}\left(\mathbb{R}\right)\right)$, and this will complete the induction. Note that, by Lemma 3.11 we know that, for each n, $T^n 0 \in C^m$. Fix $r \neq 0$, $\omega \in S^{n_0-1}$.

Define $\gamma_n = T^n\left(0\right)$, $\gamma_\infty = A$. By Lemma 3.11, we have, for $n < \infty$,

$$\begin{aligned}
\partial_u^\alpha T\left(\gamma_n\right)\left(r\omega\right) &= \frac{1}{r^{m+1}} \int_0^r s^m \partial_u^\alpha \left(-\gamma_n^2 - C_u \gamma_n - C_u\right) ds \\
&= \sum_{\alpha_1+\alpha_2=\alpha} \frac{1}{r^{m+1}} \int_0^r s^m \left(-\left(\partial_u^{\alpha_1}\gamma_n\right)\left(\partial_u^{\alpha_2}\gamma_n\right) - \left(\partial_u^{\alpha_1} C_u\right)\left(\partial_u^{\alpha_2}\gamma_n\right)\right) ds \\
&\quad - \frac{1}{r^{m+1}} \int_0^r s^m \partial_u^\alpha C_u ds.
\end{aligned} \tag{3.15}$$

Define, for $l \in C^0\left(B_{n_0}\left(\eta_1\right); \mathbb{M}_{n_0\times n_0}\left(\mathbb{R}\right)\right)$, and for $0 \leq n \leq \infty$,

$$\begin{aligned}
Q_n\left(l\right)\left(r\omega\right) = &- \sum_{\substack{\alpha_1+\alpha_2=\alpha \\ \alpha_1\neq 0 \\ \alpha_2 \neq 0}} \frac{1}{r^{m+1}} \int_0^r s^m \left(\partial_u^{\alpha_1}\gamma_n\right)\left(\partial_u^{\alpha_2}\gamma_n\right) ds \\
&- \sum_{\substack{\alpha_1+\alpha_2=\alpha \\ \alpha_1\neq 0}} \frac{1}{r^{m+1}} \int_0^r s^m \left(\partial_u^{\alpha_1} C_u\right)\left(\partial_u^{\alpha_2}\gamma_n\right) ds \\
&- \frac{1}{r^{m+1}} \int_0^r s^m \partial_u^\alpha C_u ds - \frac{1}{r^{m+1}} \int_0^r s^m \left(l\gamma_n + \gamma_n l + C_u l\right) ds.
\end{aligned} \tag{3.16}$$

Note that $Q_n\left(l\right)\left(u\right)$ extends continuously to $u=0$ and we have:

$$Q_n : C^0\left(B_{n_0}\left(\eta_1\right); \mathbb{M}_{n_0\times n_0}\left(\mathbb{R}\right)\right) \to C^0\left(B_{n_0}\left(\eta_1\right); \mathbb{M}_{n_0\times n_0}\left(\mathbb{R}\right)\right).$$

Putting (3.15) and (3.16) together, we see, for $0 \leq n < \infty$,

$$Q_n\left(\partial_u^\alpha \gamma_n\right) = \partial_u^\alpha T\left(\gamma_n\right). \tag{3.17}$$

Our next goal is to show that Q_n is a contraction ($n \leq \infty$), as a map

$$Q_n : C^0\left(B_{n_0}\left(\eta_1\right); \mathbb{M}_{n_0\times n_0}\left(\mathbb{R}\right)\right) \rightarrow C^0\left(B_{n_0}\left(\eta_1\right); \mathbb{M}_{n_0\times n_0}\left(\mathbb{R}\right)\right).$$

Consider, for $r \neq 0$, $\omega \in S^{n_0-1}$, and using that $\gamma_n \in M$ for all n,

$$\begin{aligned}
&\left\|Q_n\left(l_1\right)\left(r\omega\right) - Q_n\left(l_2\right)\left(r\omega\right)\right\| = \\
&\quad = \left\|\frac{1}{r^{m+1}}\int_0^r s^m\left[\left(l_1-l_2\right)\gamma_n + \gamma_n\left(l_1-l_2\right) + C_u\left(l_1-l_2\right)\right]\right\| \\
&\quad \leq \left\|l_1-l_2\right\|_{C^0\left(B_{n_0}(\eta_1);\mathbb{M}_{n_0\times n_0}(\mathbb{R})\right)} \frac{1}{r^{m+1}}\int_0^r s^m\left(2\kappa + Ds\right) \\
&\quad \leq \left\|l_1-l_2\right\|_{C^0\left(B_{n_0}(\eta_1);\mathbb{M}_{n_0\times n_0}(\mathbb{R})\right)}\left(\frac{2\kappa}{m+1} + \frac{D\eta_1}{m+2}\right) \\
&\quad \leq \frac{3}{4}\left\|l_1-l_2\right\|_{C^0\left(B_{n_0}(\eta_1);\mathbb{M}_{n_0\times n_0}(\mathbb{R})\right)}
\end{aligned}$$

where the last line follows by our choice of η_1.

Next, fix $l \in C^0\left(B_{n_0}\left(\eta_1\right); \mathbb{M}_{n_0\times n_0}\left(\mathbb{R}\right)\right)$. We wish to show that $Q_n\left(l\right) \rightarrow Q_\infty\left(l\right)$ in $C^0\left(B_{n_0}\left(\eta_1\right); \mathbb{M}_{n_0\times n_0}\left(\mathbb{R}\right)\right)$. Consider,

$$\begin{aligned}
Q_n\left(l\right)\left(r\omega\right) - Q_\infty\left(l\right)\left(r\omega\right) = &-\sum_{\substack{\alpha_1+\alpha_2=\alpha\\ \alpha_1\neq 0\\ \alpha_2\neq 0}} \frac{1}{r^{m+1}}\int_0^r s^m\left(\partial_u^{\alpha_1}\left(\gamma_n-\gamma_\infty\right)\right)\left(\partial_u^{\alpha_2}\gamma_n\right)ds \\
&-\sum_{\substack{\alpha_1+\alpha_2=\alpha\\ \alpha_1\neq 0\\ \alpha_2\neq 0}} \frac{1}{r^{m+1}}\int_0^r s^m\left(\partial_u^{\alpha_1}\gamma_\infty\right)\left(\partial_u^{\alpha_2}\left(\gamma_n-\gamma_\infty\right)\right)ds \\
&-\sum_{\substack{\alpha_1+\alpha_2=\alpha\\ \alpha_1\neq 0}} \frac{1}{r^{m+1}}\int_0^r s^m\left(\partial_u^{\alpha_1}C_u\right)\left(\partial_u^{\alpha_2}\left(\gamma_n-\gamma_\infty\right)\right)ds \\
&-\frac{1}{r^{m+1}}\int_0^r s^m\left(l\left(\gamma_n-\gamma_\infty\right) + \left(\gamma_n-\gamma_\infty\right)l + C_u l\right)ds.
\end{aligned}$$

Using our inductive hypothesis that $\gamma_n \rightarrow \gamma_\infty$ in $C^{m-1}\left(B_{n_0}\left(\eta_1\right); \mathbb{M}_{n_0\times n_0}\left(\mathbb{R}\right)\right)$ it is easy to show that the above goes to 0 uniformly in (r,ω) as $n \rightarrow \infty$.

In particular, if we let l_∞ be the unique fixed point of the strict contraction Q_∞ we have that $Q_n\left(l_\infty\right) \rightarrow l_\infty$ in $C^0\left(B_{n_0}\left(\eta_1\right); \mathbb{M}_{n_0\times n_0}\left(\mathbb{R}\right)\right)$. Using (3.17), we have:

$$\partial_u^\alpha \gamma_{n+1} = \partial_u^\alpha T\left(\gamma_n\right) = Q_n\left(\partial_u^\alpha \gamma_n\right).$$

Hence, Lemma 3.12 shows that:

$$\partial_u^\alpha \gamma_n \rightarrow l_\infty$$

which shows that $\partial_u^\alpha T^n 0$ converges in $C^0\left(B_{n_0}\left(\eta_1\right); \mathbb{M}_{n_0\times n_0}\left(\mathbb{R}\right)\right)$. It follows that $A \in C^m\left(B_{n_0}\left(\eta_1\right); \mathbb{M}_{n_0\times n_0}\left(\mathbb{R}\right)\right)$.

Moreover, we have that $\partial_u^\alpha A = l_\infty$, where l_∞ was the unique fixed point of Q_∞. Hence, by the contraction mapping principle, $\partial_u^\alpha A = \lim_{n\to\infty} Q_\infty^n 0$. It follows, by a proof similar to the one we did before for T, that we have:

$$\left\|\partial_u^\alpha A\right\|_{C^0\left(B_{n_0}(\eta_1);\mathbb{M}_{n_0\times n_0}(\mathbb{R})\right)} \leq 4\left\|Q_\infty\left(0\right)\right\|_{C^0\left(B_{n_0}(\eta_1);\mathbb{M}_{n_0\times n_0}(\mathbb{R})\right)}.$$

Let us suppose, for induction that we have (3.11) for $m-1$. Then to prove (3.11) for m it suffices to show that:

$$\left\|Q_\infty\left(0\right)\right\|_{C^0\left(B_{n_0}(\eta_1);\mathbb{M}_{n_0\times n_0}(\mathbb{R})\right)} \leq C_m\left(m, \widetilde{C}_{m,\eta_1}\right)$$

but this follows immediately from the inductive hypothesis and the definition of Q_∞. ■

Now fix η_1 and A as in the conclusion of Theorem 3.10, taking $\kappa = \frac{1}{2}$.

Lemma 3.13. $A\big|_{B_{n_0}(\delta)} = \widehat{A}$.

Proof. Using, as remarked before, that for fixed ω, $\left\|\widehat{A}\left(r\omega\right)\right\| = O\left(r\right)$, this follows just as in the proof of uniqueness in Theorem 3.10. ■

Lemma 3.13 shows that we may extend the vector fields $\widehat{Y}_j$ by setting:

$$Y_j = \partial_{u_j} + \sum_{k=1}^{n_0} a_j^k \partial_{u_k}$$

where $A_{j,k} = \left(a_j^k\right)$.

Theorem 3.14. $d\Phi\left(Y_j\right) = X_j$.

To prove Theorem 3.14, we need a preliminary lemma:

Lemma 3.15. *Fix* $\omega \in S^{n_0-1}$, $r_0 < \eta_1$, *and suppose that for all* $r \leq r_0$, $\left|\det_{n_0\times n_0} d\Phi\left(r\omega\right)\right| \neq 0$. *Then, on the line* $\left\{r\omega : 0 \leq r \leq r_0\right\}$, *we have* $d\Phi\left(Y_j\right) = X_j$, $1 \leq j \leq n_0$.

Proof. Suppose not. Define

$$r_1 = \sup\left\{r \geq 0 : d\Phi\left(Y_j\right) = X_j \text{ on the line } \left\{r'\omega : 0 \leq r' \leq r\right\}, 1 \leq j \leq n_0\right\}.$$

Then we must have $r_1 < r_0$ (by continuity). Since $\widehat{Y}_j = Y_j\big|_{B_{n_0}(\delta)}$, we know that $r_1 > 0$. Since $\left|\det_{n_0\times n_0} d\Phi\left(r_1\omega\right)\right| \neq 0$, the inverse function theorem implies that there exists a neighborhood V of $r_1\omega$ such that $\Phi : V \to \Phi\left(V\right)$ is a C^1 diffeomorphism.

Pick $0 < r_2 < r_3 < r_1 < r_4$ such that:

$$\{r'\omega : r_2 \leq r' \leq r_4\} \subset V.$$

Let $\widetilde{Y}_j$ be the pullback of X_j to V via the map Φ. By our choice of r_1, we have that on the line $\{r'\omega : r_2 \leq r' \leq r_3\}$, $\widetilde{Y}_j = Y_j$. On the other hand, if we write:

$$\widetilde{Y}_j = \partial_{u_j} + \tilde{a}_j^k \partial_{u_k}$$

then the coefficients $\tilde{a}_j^k$ satisfy the differential equation (3.10) (this follows just as before). Away from $r = 0$ this is a standard ODE, so standard uniqueness theorems (say using Gronwall's inequality) show that $\widetilde{Y}_j = Y_j$ on the line $\{r'\omega : r_2 \leq r' \leq r_4\}$. This contradicts our choice of r_1. ■

From here, Theorem 3.14 will follow immediately from the following theorem:

Theorem 3.16. *For all* $t \in B_{n_0}(\eta_1)$*,*

$$\left|\det_{n_0 \times n_0} d\Phi(t)\right| \approx \left|\det_{n_0 \times n_0} d\Phi(0)\right| = \left|\det_{n_0 \times n_0} X(x_0)\right|.$$

Theorem 3.16, in turn, follows immediately from the following lemma and a simple continuity argument:

Lemma 3.17. *Fix* $\omega \in S^{n_0 - 1}$, $0 < r_0 < \eta_1$*. Suppose for* $0 \leq r \leq r_0$, $|\det_{n_0 \times n_0} d\Phi(r\omega)| \neq 0$*. Then, for all* $0 \leq r \leq r_0$*,*

$$\left|\det_{n_0 \times n_0} d\Phi(r\omega)\right| \approx \left|\det_{n_0 \times n_0} d\Phi(0)\right|$$

Here, the implicit constants depend on neither r_0 *nor* ω*.*

Proof. By Lemma 3.15, for all $0 \leq r \leq r_0$, we have

$$d\Phi(Y_j)(\Phi(r\omega)) = X_j(\Phi(r\omega)).$$

Rewriting this in matrix notation, we have:

$$d\Phi((I + A) \bigtriangledown_u)(\Phi(r\omega)) = X(\Phi(r\omega)).$$

Thus, by applying (B.1), we have at the point $\Phi(r\omega)$:

$$\left|\det_{n_0 \times n_0} X\right| = \left|\det_{n_0 \times n_0} d\Phi(I + A)\right| = \sqrt{\det\left((I + A)^t d\Phi^t d\Phi(I + A)\right)}$$
$$= |\det(I + A)| \sqrt{\det(d\Phi^t d\Phi)} = |\det(I + A)| \left|\det_{n_0 \times n_0} d\Phi\right|.$$

However, we have $\|A\| \leq \frac{1}{2}$, and so we have:

$$\left|\det_{n_0\times n_0} d\Phi\left(r\omega\right)\right| \approx \left|\det_{n_0\times n_0} X\left(\Phi\left(r\omega\right)\right)\right|.$$

Applying Lemma 3.8, we see that

$$\left|\det_{n_0\times n_0} X\left(\Phi\left(r\omega\right)\right)\right| \approx \left|\det_{n_0\times n_0} X\left(x_0\right)\right| = \left|\det_{n_0\times n_0} d\Phi\left(0\right)\right|,$$

completing the proof. ∎

Proposition 3.18. *We have, for $m \geq 0$:*

$$\|f\|_{C^{m+1}\left(B_{n_0}(\eta_1)\right)} \approx_m \sum_{|\alpha|\leq m+1} \|Y^\alpha f\|_{C^0\left(B_{n_0}(\eta)\right)}$$

and,

$$\|A\|_{C^m\left(B_{n_0}(\eta_1);\mathbb{M}_{n_0\times n_0}(R)\right)} \lesssim_m 1.$$

It immediately follows that:

$$\|Y_j\|_{C^m\left(B_{n_0}(\eta_1)\right)} \lesssim_m 1$$

in particular,

$$\|Y_j\|_{C^2\left(B_{n_0}(\eta_1)\right)} \lesssim 1.$$

Proof. We prove the result by induction. Our base case will be $m = 0$. We already know,

$$\|A\|_{C^0\left(B_{n_0}(\eta_1)\right)} \leq \frac{1}{2} \lesssim_0 1.$$

Recall, $\lesssim_0, \lesssim_1, \lesssim_2$, and $\lesssim$ all mean the same thing. For notational convenience, write the operator:

$$\bigtriangledown_Y f = \left(Y_1 f, \ldots, Y_{n_0} f\right).$$

Since $\bigtriangledown_Y = (I + A) \bigtriangledown_u$, and $\|I + A\|_{C^0\left(B_{n_0}(\eta_1);\mathbb{M}_{n_0\times n_0}(\mathbb{R})\right)} \lesssim 1$, it follows that:

$$\sum_{|\alpha|\leq 1} \|Y^\alpha f\|_{C^0\left(B_{n_0}(\eta_1)\right)} \lesssim_0 \|f\|_{C^1\left(B_{n_0}(\eta_1)\right)}.$$

Conversely, since $\bigtriangledown_u = (I+A)^{-1} \bigtriangledown_Y$ and $\left\|(I + A)^{-1}\right\|_{C^0\left(B_{n_0}(\eta_1);\mathbb{M}_{n_0\times n_0}(\mathbb{R})\right)} \lesssim 1$ (which can be seen by writing $(I + A)^{-1}$ as a Neumann series), we have:

$$\|f\|_{C^1\left(B_{n_0}(\eta_1)\right)} \lesssim_0 \sum_{|\alpha|\leq 1} \|Y^\alpha f\|_{C^0\left(B_{n_0}(\eta_1)\right)}.$$

Suppose, for induction, that we have:

$$\|f\|_{C^m\left(B_{n_0}(\eta_1)\right)} \approx_{m-1} \sum_{|\alpha|\leq m} \|Y^\alpha f\|_{C^0\left(B_{n_0}(\eta)\right)}.$$

Then, note,

$$\left\|\tilde{c}_{i,j}^k\right\|_{C^m\left(B_{n_0}(\eta_1)\right)} \approx_{m-1} \sum_{|\alpha|\leq m} \left\|Y^\alpha \tilde{c}_{i,j}^k\right\|_{C^0\left(B_{n_0}(\eta)\right)}.$$

But,

$$Y^\alpha \tilde{c}_{i,j}^k = \left(X^\alpha c_{i,j}^k\right) \circ \Phi$$

and hence,

$$\sum_{|\alpha|\leq m} \left\|Y^\alpha \tilde{c}_{i,j}^k\right\|_{C^0\left(B_{n_0}(\eta)\right)} \lesssim_m 1$$

and we have that:

$$\left\|\tilde{c}_{i,j}^k\right\|_{C^m\left(B_{n_0}(\eta_1)\right)} \lesssim_m 1$$

for all i, j, k. It follows from Theorem 3.10 that:

$$\|A\|_{C^m\left(B_{n_0}(\eta_1);\mathbb{M}_{n_0\times n_0}(R)\right)} \lesssim_m 1.$$

And thus, we have, using the Neumann series, that:

$$\left\|(I+A)^{-1}\right\|_{C^m\left(B_{n_0}(\eta_1);\mathbb{M}_{n_0\times n_0}(\mathbb{R})\right)}, \quad \|I+A\|_{C^m\left(B_{n_0}(\eta_1);\mathbb{M}_{n_0\times n_0}(\mathbb{R})\right)} \lesssim_m 1.$$

Hence, since $\bigtriangledown_u = (I+A)^{-1} \bigtriangledown_Y$ and $\bigtriangledown_Y = (I+A) \bigtriangledown_u$, it follows easily that:

$$\|f\|_{C^{m+1}\left(B_{n_0}(\eta_1)\right)} \approx_m \sum_{|\alpha|\leq m+1} \|Y^\alpha f\|_{C^0\left(B_{n_0}(\eta)\right)}.$$

■

Now we turn our attention to showing that if we shrink η_1 enough, while still keeping it admissible, we have that Φ is injective on $B_{n_0}(\eta_1)$. This result is essentially contained in [30] (see p. 622 of that reference), however we recreate the proof below for completeness, and to make it clear why each constant is admissible. Thus, the next lemma and proposition follow [30].

Lemma 3.19. *Suppose Z is a C^1 vector field on an open subset $V \subseteq \mathbb{R}^n$, and $U \subseteq V$. Then, there exists a $\delta > 0$, depending only on n, such that if $\|Z\|_{C^1(U)} \leq \delta$, then there does not exist $x_1 \in U$ with:*

- $e^{tZ} x_1 \in U$, $0 \leq t \leq 1$,
- $e^Z x_1 = x_1$,
- $Z(x_1) \neq 0$.

Proof. Suppose the lemma does not hold, and we have an x_1 and Z as above. In the proof of this lemma, we will use big-O notation–the implicit constants will only depend on n. Differentiating the identity:

$$\frac{d}{dt}e^{tZ}x_1 = Z\left(e^{tZ}x_1\right)$$

we obtain:

$$\frac{d^2}{dt^2}e^{tZ}x_1 = O\Big(\delta\Big|\frac{d}{dt}e^{tZ}x_1\Big|\Big).$$

Thus, by Gronwall's inequality:

$$\frac{d}{dt}e^{tZ}x_1 = O\Big(\Big|\frac{d}{dt}e^{tZ}x_1\Big|_{t=0}\Big|\Big) = O(|Z(x_1)|)$$

for $t \leq 1$. Hence,

$$\frac{d^2}{dt^2}e^{tZ}x_1 = O\left(\delta\left|Z\left(x_1\right)\right|\right).$$

Integrating, we obtain:

$$\frac{d}{dt}e^{tZ}x_1 = Z\left(x_1\right) + O\left(\delta\left|t\right|\left|Z\left(x_1\right)\right|\right).$$

Integrating again, we obtain:

$$x_1 = e^{Z}x_1 = x_1 + Z\left(x_1\right) + O\left(\delta\left|Z\left(x_1\right)\right|\right).$$

which is impossible if δ is sufficiently small, completing the proof. ■

Proposition 3.20. *We may shrink η_1, while still keeping it admissible, to ensure that Φ is injective on $B_{n_0}(\eta_1)$.*

Proof. We will construct an admissible constant η_2 with the properties desired in the statement of the proposition, and then the proof will be completed by renaming η_2, η_1.

Consider the maps $\Psi_{u_0}(u) = e^{u\cdot Y}u_0$, defined for $|u_0|, |u| \leq \eta'$, where $\eta' > 0$ is some sufficiently small admissible constant. Notice, since

$$\|Y_j\|_{C^2\left(B_{n_0}(\eta_1)\right)} \lesssim 1$$

we have by Theorem A.1 that $\Psi_{u_0} \in C^2$ with C^2 norm admissibly bounded uniformly in u_0. Furthermore, since $d\Psi_{u_0}(0) = (I + A(u_0))$, and $\|A\| \leq \frac{1}{2}$, we have that $|\det d\Psi_{u_0}(0)| \gtrsim 1$, uniformly in u_0.

Hence we apply the uniform inverse function theorem (Theorem A.3) to see that there exist admissible constants $\eta_2 > 0$, $\delta > 0$ such that for all $u_1, u_2 \in B_{n_0}(\eta_2)$ there exists $u_0 \in B_{n_0}(\delta)$ with $u_2 = \Psi_{u_1}(u_0)$. Moreover, by shrinking η_2, we may shrink δ.

Now suppose $\Phi : B(\eta_2) \to \widetilde{B}_{(X,d)}(x_0,\xi)$ is not injective. Thus, there exist $u_1, u_2 \in B_{n_0}(\eta_2), u_1 \neq u_2$ such that

$$\Phi(u_1) = \Phi(u_2).$$

But, since there exists $0 \neq u_0 \in B_{n_0}(\delta)$ with $u_2 = e^{u_0 \cdot Y} u_1$, we have that:

$$\Phi(u_1) = \Phi(u_2) = e^{u_0 \cdot X}\Phi(u_1).$$

Setting $Z = u_0 \cdot X$, we have by Lemma 3.8 that Z is non-zero on $B_{(X,d)}(x_0,\xi)$. Applying Lemma 3.19, we see that by taking δ admissibly small enough (and therefore η_2 admissibly small enough), we achieve a contradiction. ∎

Our proof of Theorem 3.1 will now be completed by the following proposition:

Proposition 3.21. ([22, Lemma 2.16]) *There exists an admissible constant* $\xi_1 > 0$, *such that*

$$B_{(X,d)}(x_0,\xi_1) \subseteq \Phi(B_{n_0}(\eta_1)).$$

Proof. Actually, the proof in [22] proves something more general. In our case, though, we have already shown that Φ is injective (Proposition 3.20), and this simplifies matters, somewhat. We include this simplified proof, and refer the reader to [22] for the stronger results.

Fix $\xi_1 > 0$. Suppose $y \in B_{(X,d)}(x_0,\xi_1)$. Thus, there exists $\phi : [0,1] \to B_{(X,d)}(x_0,\xi_1)$, $\phi(0) = x_0$, $\phi(1) = y$,

$$\phi'(t) = (b \cdot X)(\phi(t))$$

with $b \in L^\infty([0,1])^{n_0}$, $\left\|\left|\xi_1^{-d} b\right|\right\|_{L^\infty([0,1])} < 1$.

Define

$$\mathcal{T} = \left\{ t \leq 1 : \phi(t') \in \Phi\left(B_{n_0}\left(\frac{\eta_1}{2}\right)\right), \forall 0 \leq t' \leq t \right\}.$$

Let $t_0 = \sup \mathcal{T}$. We want to show that, by taking ξ_1 admissibly small enough we have that $t_0 = 1$ and $\phi(1) \in \Phi\left(B_{n_0}\left(\frac{\eta_1}{2}\right)\right)$.

Suppose not. Then, we must have that $\left|\Phi^{-1}(\phi(t_0))\right| = \frac{\eta_1}{2}$. Then, we have:

$$\begin{aligned}\frac{\eta_1}{2} &= \left|\Phi^{-1}(\phi(t_0))\right| = \left|\int_0^{t_0} \frac{d}{dt}\Phi^{-1}(\phi(t))\right| \\ &= \left|\int_0^{t_0} \left((b \cdot X)\Phi^{-1}\right)(\phi(t))\right| = \left|\int_0^{t_0} (b \cdot Y)\left(\Phi^{-1}(\phi(t))\right)\right| < \frac{\eta_1}{2},\end{aligned}$$

provided ξ_1 is admissibly small enough. In the second to last line,

$$(b \cdot Y)\left(\Phi^{-1}(\phi(t))\right)$$

denotes the vector $b \cdot Y$ evaluated at the point $\Phi^{-1}(\phi(t))$. This achieves the contradiction and completes the proof. ∎

4. Carnot-Carathéodory balls at the unit scale

In this section, we generalize Theorem 3.1 to the case when the vector fields may not be linearly independent–thereby completing the proof of Theorem 1.3. Suppose $X = (X_1, \dots, X_q)$ are q C^1 vector fields with associated single-parameter formal degrees $d = (d_1, \dots, d_q) \in (0, \infty)^q$, defined on the fixed connected open set $\Omega \subseteq \mathbb{R}^n$. Fix $1 \geq \xi > 0$, $x_0 \in \Omega$. Let $n_0 = \dim \operatorname{span}\{X_1(x_0), \dots, X_q(x_0)\}$. Fix $1 \geq \zeta > 0$, $J_0 \in \mathcal{I}(n_0, q)$; we assume that:

$$\left| \det_{n_0 \times n_0} X_{J_0}(x_0) \right|_\infty \geq \zeta \sup_{J \in \mathcal{I}(n_0,q)} \left| \det_{n_0 \times n_0} X_J(x_0) \right|_\infty . \tag{4.1}$$

Recall if $J_0 = (j_1, \dots, j_{n_0})$, $(X, d)_{J_0}$ denotes the list with formal degrees $\left((X_{j_1}, d_{j_1}), \dots, (X_{j_{n_0}}, d_{j_{n_0}})\right)$. We also write X_{J_0} to denote the list of vector field $X_{j_1}, \dots, X_{j_{n_0}}$, and d_{J_0} to denote the list of formal degrees $d_{j_1}, \dots, d_{j_{n_0}}$. Suppose, further, that $(X, d)_{J_0}$ satisfies $\mathcal{C}(x_0, \xi)$. In addition, suppose that the X_js satisfy an integrability condition on $B_{(X,d)_{J_0}}(x_0, \xi)$ given by:

$$[X_j, X_k] = \sum_l c_{j,k}^l X_l. \tag{4.2}$$

Without loss of generality, we assume for the remainder of the section that $J_0 = (1, \dots, n_0)$. We will also assume that:

- For $1 \leq j \leq n_0$, X_j is C^2 on $B_{(X,d)_{J_0}}(x_0, \xi)$ and satisfies
$$\|X_j\|_{C^2\left(B_{(X,d)_{J_0}}(x_0,\xi)\right)} < \infty .$$
- For $|\alpha| \leq 2$, $1 \leq i, j \leq n_0$, $1 \leq k \leq q$, $X_{J_0}^\alpha c_{i,j}^k \in C^0\left(B_{(X,d)}(x_0, \xi)\right)$, and
$$\sum_{|\alpha| \leq 2} \left\| X_{J_0}^\alpha c_{i,j}^k \right\|_{C^0\left(B_{(X,d)_{J_0}}(x_0,\xi)\right)} < \infty.$$
- For $|\alpha| \leq 1$, $1 \leq i, j, k \leq q$, $X_{J_0}^\alpha c_{i,j}^k \in C^0\left(B_{(X,d)}(x_0, \xi)\right)$, and
$$\sum_{|\alpha| \leq 1} \left\| X_{J_0}^\alpha c_{i,j}^k \right\|_{C^0\left(B_{(X,d)_{J_0}}(x_0,\xi)\right)} < \infty.$$

We will say that C is an admissible constant if C can be chosen to depend only on a fixed upper bound, $d_{max} < \infty$, for $d_1, \dots, d_q$, a fixed lower bound $d_{min} > 0$ for $d_1, \dots, d_q$, a fixed upper bound for n and q (and therefore for n_0), a fixed lower bound, $\xi_0 > 0$, for ξ, a fixed lower bound, $\zeta_0 > 0$, for ζ, and a fixed upper bound for the quantities:

$$\begin{aligned} &\left\|X_j\right\|_{C^2\left(B_{(X,d)_{J_0}}(x_0,\xi)\right)}, \quad 1 \leq j \leq n_0, \\ \sum_{|\alpha|\leq 2} &\left\|X_{J_0}^\alpha c_{i,j}^k\right\|_{C^0\left(B_{(X,d)_{J_0}}(x_0,\xi)\right)}, \quad 1 \leq i,j \leq n_0, \quad 1 \leq k \leq q, \\ \sum_{|\alpha|\leq 1} &\left\|X_{J_0}^\alpha c_{i,j}^k\right\|_{C^0\left(B_{(X,d)_{J_0}}(x_0,\xi)\right)}, \quad 1 \leq i,j,k \leq q. \end{aligned}$$

Furthermore, if we say that C is an m-admissible constant, we mean that in addition to the above, we assume that:

- $\left\|X_j\right\|_{C^m\left(B_{(X,d)_{J_0}}(x_0,\xi)\right)} < \infty$, for every $1 \leq j \leq n_0$,
- $\sum\limits_{|\alpha|\leq m} \left\|X_{J_0}^\alpha c_{i,j}^k\right\|_{C^0\left(B_{(X,d)_{J_0}}(x_0,\xi)\right)} < \infty$, for every $1 \leq i,j \leq n_0$, $1 \leq k \leq q$,
- $\sum_{|\alpha|\leq m-1} \left\|X_{J_0}^\alpha c_{i,j}^k\right\|_{C^0\left(B_{(X,d)_{J_0}}(x_0,\xi)\right)} < \infty$, for every $1 \leq i,j,k \leq q$.

(in particular, the above partial derivatives exist and are continuous). C is allowed to depend on m, all the quantities an admissible constant is allowed to depend on, and a fixed upper bound for the above quantities. Note that, as before, $\lesssim_0, \lesssim_1, \lesssim_2$, and $\lesssim$ all denote the same thing.

For $\eta > 0$, a sufficiently small admissible constant, define the map:

$$\Phi : B_{n_0}(\eta) \to \widetilde{B}_{(X,d)_{J_0}}(x_0, \xi)$$

by

$$\Phi(u) = \exp(u \cdot X_{J_0}) x_0.$$

The main results of this section are the following:

Theorem 4.1. *There exist admissible constants $\eta_1 > 0$, $\xi_1 \geq \xi_2 > 0$, such that:*

- $\Phi : B_{n_0}(\eta_1) \to \widetilde{B}_{(X,d)_{J_0}}(x_0, \xi)$ *is one-to-one.*
- *For all $u \in B_{n_0}(\eta_1)$, $\left|\det_{n_0\times n_0} d\Phi(u)\right| \approx \left|\det_{n_0\times n_0} X(x_0)\right|$.*
- $B_{(X,d)}(x_0,\xi_2) \subseteq B_{(X,d)_{J_0}}(x_0,\xi_1) \subseteq \Phi(B_{n_0}(\eta_1)) \subseteq \widetilde{B}_{(X,d)_{J_0}}(x_0,\xi) \subseteq B_{(X,d)_{J_0}}(x_0,\xi) \subseteq B_{(X,d)}(x_0,\xi)$.

Furthermore, if we let Y_j *($1 \leq j \leq q$) be the pullback of* X_j *under the map* Φ*, then we have:*

$$\|Y_j\|_{C^m\left(B_{n_0}(\eta_1)\right)} \lesssim_m 1$$

in particular,

$$\|Y_j\|_{C^2\left(B_{n_0}(\eta_1)\right)} \lesssim 1.$$

Finally, if for $u \in B_{n_0}(\eta_1)$ *we define the* $n_0 \times n_0$ *matrix* $A(u)$ *by:*[21]

$$(Y_1, \ldots, Y_{n_0}) = (I + A) \bigtriangledown_u$$

then,

$$\sup_{u \in B_{n_0}(\eta_1)} \|A(u)\| \leq \frac{1}{2}.$$

Corollary 4.2. *Let* η_1, ξ_1, ξ_2 *be as in Theorem* 4.1*. Then, there exist admissible constants* $0 < \eta_2 < \eta_1$*,* $0 < \xi_4 \leq \xi_3 < \xi_2$ *such that:*

$$\begin{aligned} B_{(X,d)}(x_0, \xi_4) &\subseteq B_{(X,d)_{J_0}}(x_0, \xi_3) \subseteq \Phi(B_{n_0}(\eta_2)) \\ &\subseteq \widetilde{B}_{(X,d)_{J_0}}(x_0, \xi_2) \subseteq B_{(X,d)_{J_0}}(x_0, \xi_2) \subseteq B_{(X,d)}(x_0, \xi_2) \\ &\subseteq B_{(X,d)_{J_0}}(x_0, \xi_1) \subseteq \Phi(B_{n_0}(\eta_1)) \subseteq \widetilde{B}_{(X,d)_{J_0}}(x_0, \xi) \\ &\subseteq B_{(X,d)_{J_0}}(x_0, \xi) \subseteq B_{(X,d)}(x_0, \xi), \end{aligned}$$

and $\mathrm{Vol}\left(B_{(X,d)}(x_0, \xi_2)\right) \approx |\det_{n_0 \times n_0} X(x_0)|$*, where* $\mathrm{Vol}(A)$ *denotes the induced Lebesgue volume on the leaf generated by the* X_j*s, passing through the point* x_0*.*

Corollary 4.3. *Take* ξ_4 *as in Corollary* 4.2*. There exists* $\phi \in C_0^2(B_{(X,d)}(x_0, \xi))$ *(here, we mean* C^2 *as thought of as a function on the leaf), which equals* 1 *on* $B_{(X,d)}(x_0, \xi_4)$ *and satisfies:*

$$|X^\alpha \phi| \lesssim_{(|\alpha|-1)\vee 0} 1$$

for every ordered multi-index α*.*

Remark 4.4. Later in the paper we will apply Corollaries 4.2 and 4.3 without explicitly saying what J_0 and ζ are. In these cases, we are choosing $\zeta = 1$ and J_0 such that:

$$\left| \det_{n_0 \times n_0} X_{J_0}(x_0) \right|_\infty = \left| \det_{n_0 \times n_0} X(x_0) \right|_\infty.$$

[21] Recall, we have, without loss of generality, assumed $J_0 = (1, \ldots, n_0)$.

Remark 4.5. In our definition of admissible constants, we have assumed greater regularity on $X_1, \ldots, X_{n_0}$ than on X_j, $n_0 < j \leq q$. In many applications, it is easier to just assume more symmetric regularity assumptions, that imply the assumptions of this section. Later in the paper we sometimes assume the following, stronger hypotheses:

- (X, d) satisfies $\mathcal{C}(x_0, \xi)$, and (4.2) holds on $B_{(X,d)}(x_0, \xi)$.
- In addition to everything that they are allowed to depend on in this section, m-admissible constants (for $m \geq 2$) can depend on a fixed upper bound for the quantities:

$$\left\|X_l\right\|_{C^m\left(B_{(X,d)}(x_0,\xi)\right)}, \quad \sum_{|\alpha| \leq m} \left\|X^\alpha c_{i,j}^k\right\|_{C^0\left(B_{(X,d)}(x_0,\xi)\right)}$$

  where $1 \leq i, j, k, l \leq q$, and these derivatives are assumed to exist, and the norms are assumed to be finite. Admissible constants are defined to be 2-admissible constants.

Remark 4.6. Just as in Remark 3.3, the d_js do not play an essential role in this section.

Remark 4.7. As mentioned in the Section 1.1, "at the unit scale" in the title of this section refers to the unit scale with respect to the vector fields X_j. Thus, if the vector fields X_j are very small, one can think of the results in this section as taking place at a very small scale.

Remark 4.8. The observant reader may have noticed that we made no *a priori* bound on X_j, $n_0 < j \leq q$. However, we will see using Cramer's rule that (4.1) implies a bound for X_j at x_0. In addition, we will be able to use Gronwall's inequality to obtain bounds at points other than x_0 (see (4.4)). The reader may wonder, though, that since we have assumed no *a priori* bound for the C^1 norm of X_j ($n_0 < j \leq q$), do we need to insist that they are C^1? The answer is partially no, though our definitions only makes sense when the X_j are all assumed to be C^1. We will see that the above assumptions will show that $X_1, \ldots, X_{n_0}$ are integrable (see Proposition 4.14), and from there all we need is that X_j ($n_0 < j \leq q$) is C^1 on the leaf generated by $X_1, \ldots, X_{n_0}$, passing through x_0. This perspective is taken up in Section 4.1; in fact, the main reason we have been careful to *not* assume a bound on the C^1 norm of X_j ($n_0 < j \leq q$) in this section, is to make clear how these arguments also work in the setup of Section 4.1.

Before we prove Theorem 4.1, let us first see how it implies the two corollaries.

Proof of Corollary 4.2. We obtain η_1, ξ_1, ξ_2 from Theorem 4.1. Then, apply Theorem 4.1 again with ξ_2 in place of ξ to complete the proof of the first part of the corollary.

By the above containments, we have:

$$\mathrm{Vol}\left(\Phi\left(B_{n_0}\left(\eta_2\right)\right)\right) \lesssim \mathrm{Vol}\left(B_{(X,d)}\left(x_0, \xi_2\right)\right) \lesssim \mathrm{Vol}\left(\Phi\left(B_{n_0}\left(\eta_1\right)\right)\right).$$

Using (B.2) and the fact that

$$\left|\det_{n_0 \times n_0} d\Phi\left(t\right)\right| \approx \left|\det_{n_0 \times n_0} X\left(x_0\right)\right|$$

for all $t \in B_{n_0}\left(\eta_1\right)$, the estimate on the volume follows immediately. ■

Remark 4.9. By a proof similar to the one of Corollary 4.2, we have that if $\xi' > 0$, is a fixed admissible constant with $\xi' \leq \xi_2$, then,

$$\mathrm{Vol}\left(B_{(X,d)}\left(x_0, \xi'\right)\right) \approx \left|\det_{n_0 \times n_0} X\left(x_0\right)\right|.$$

Proof of Corollary 4.3. Let $\psi \in C_0^\infty\left(B_{n_0}\left(\eta_1\right)\right)$, with $\psi = 1$ on $B_{n_0}\left(\eta_2\right)$. Define

$$\phi\left(x\right) = \begin{cases} \psi\left(\Phi^{-1}\left(x\right)\right) & \text{if } x \in \Phi\left(B_{n_0}\left(\eta_1\right)\right), \\ 0 & \text{otherwise.} \end{cases}$$

Then, we see:

$$X^\alpha \phi\left(x\right) = \begin{cases} \left(Y^\alpha \psi\right)\left(\Phi^{-1}\left(x\right)\right) & \text{if } x \in \Phi\left(B_{n_0}\left(\eta_1\right)\right), \\ 0 & \text{otherwise.} \end{cases}$$

Thus, to prove the corollary, it suffices to show that:

$$\left|Y^\alpha \psi\right| \lesssim_{(|\alpha|-1)\vee 0} 1$$

and this is obvious. ■

We now turn to the proof of Theorem 4.1. The main idea is to apply Theorem 3.1 to the vector fields $(X,d)_{J_0}$.

Lemma 4.10. *Fix $1 \leq n_1 \leq n \wedge q$. Then, for $1 \leq j \leq n_0$, $I \in \mathcal{I}\left(n_1, n\right)$, $J \in \mathcal{I}\left(n_1, q\right)$, $x \in B_{(X,d)_{J_0}}\left(x_0, \xi\right)$,*

$$\left|X_j \det X\left(x\right)_{I,J}\right| \lesssim \left|\det_{n_1 \times n_1} X\left(x\right)\right|.$$

Proof. This can be proved by a simple modification of the proof for Lemma 3.6. We leave the details to the reader. ■

Lemma 4.11. *For* $y \in B_{(X,d)_{J_0}}(x_0,\xi)$, $1 \le n_1 \le q \wedge n$,

$$\left|\det_{n_1\times n_1} X(y)\right| \approx \left|\det_{n_1\times n_1} X(x_0)\right|.$$

In particular, for all $y \in B_{(X,d)_{J_0}}(x_0,\xi)$, $\dim \operatorname{span}\{X_1(y), \ldots X_q(y)\} = n_0$.

Proof. This can be proved by a simple modification of the proof of Lemma 3.8, using Lemma 4.10. We leave the details to the reader. ■

Take $I_0 \in \mathcal{I}(n_0,n)$ such that:

$$\left|\det X(x_0)_{I_0,J_0}\right| = \sup_{I\in\mathcal{I}(n_0,n)} \left|\det X(x_0)_{I,J_0}\right|.$$

Lemma 4.12. *There exists an admissible constant* $\xi^1 > 0$, $\xi^1 \le \xi$ *such that for every* $y \in B_{(X,d)_{J_0}}(x_0,\xi^1)$, *we have:*

$$\left|\det X(y)_{I_0,J_0}\right| \gtrsim \left|\det_{n_0\times n_0} X(y)\right|$$

Proof. Fix $I \in \mathcal{I}(n_0,n)$, $J \in \mathcal{I}(n_0,q)$. Let $\gamma : [0,1] \to B_{(X,d)_{J_0}}(x_0,\xi)$ satisfy:

$$\gamma'(t) = (b\cdot X_{J_0})(\gamma(t))$$

with $b \in L^\infty([0,1])^{n_0}$, $\left\|\left|\xi^{-d_{J_0}} b\right|\right\|_{L^\infty([0,1])} < 1$.

By applying Lemmas 4.10, 4.11, we see:

$$\begin{aligned}\frac{d}{dt}\left|\det X_{I,J}(\gamma(t))\right|^2 &= \det X_{I,J}(\gamma(t))\left((b\cdot X_{J_0})\det X_{I,J}\right)(\gamma(t))\\ &\lesssim \left|\det_{n_0\times n_0} X(\gamma(t))\right|^2 \approx \left|\det_{n_0\times n_0} X(x_0)\right|^2\\ &\approx \left|\det_{n_0\times n_0} X(x_0)_{J_0}\right|^2 \approx \left|\det X(x_0)_{I_0,J_0}\right|^2\end{aligned}$$

and therefore,

$$\frac{d}{dt}\left|\det X_{I,J}(\gamma(t))\right|^2 \le C\left|\det X(x_0)_{I_0,J_0}\right|^2$$

where C is some admissible constant. Thus, if $t \le \frac{1}{2C}$, we have:

$$\left|\det X(\gamma(t))_{I_0,J_0}\right| \approx \left|\det X(x_0)_{I_0,J_0}\right|$$

and,

$$\begin{aligned}\left|\det X(\gamma(t))_{I,J}\right| &\lesssim \left|\det X(x_0)_{I,J}\right| + \left|\det X(x_0)_{I_0,J_0}\right|\\ &\lesssim \left|\det X(x_0)_{I_0,J_0}\right| \approx \left|\det X(\gamma(t))_{I_0,J_0}\right|.\end{aligned} \tag{4.3}$$

We complete the proof by noting that there exists an admissible constant $\xi^1 > 0$ such that for every point $y \in B_{(X,d)_{J_0}}(x_0, \xi^1)$ there is a γ of the above form and a $t \leq \frac{1}{2C}$ with $y = \gamma(t)$. ■

Lemma 4.13. *Fix $I \in \mathcal{I}(n_0, n)$, $J \in \mathcal{I}(n_0, q)$. Then,*

$$\sum_{|\alpha| \leq m} \left\| X_{J_0}^{\alpha} \frac{\det X_{I,J}}{\det X_{I_0,J_0}} \right\|_{C^0\left(B_{(X,d)_{J_0}}(x_0,\xi^1)\right)} \lesssim_m 1.$$

Proof. For $m = 0$ this follows from Lemma 4.12. For $m > 0$, we look back to the proof of Lemma 3.6. There, it was shown that $X_j \det X_{I,J}$ could be written as a sum of terms of the form

$$f \det X_{I',J'}$$

where $I' \in \mathcal{I}(n_0, n)$, $J' \in \mathcal{I}(n_0, q)$, and f was either of the form $c_{i,j}^k$ or f was a derivative of a coefficient of X_j $(1 \leq j \leq n_0)$. From this, Lemma 4.12, and a simple induction, the lemma follows easily. We leave the proof to the interested reader. ■

We now show that on $B_{(X,d)_{J_0}}(x_0, \xi^1)$, the vector fields $X_1, \ldots, X_{n_0}$ satisfy the hypotheses of Theorem 3.1. Recall, we have assumed, without loss of generality, $J_0 = (1, \ldots, n_0)$.

Proposition 4.14. *For $1 \leq i, j, k \leq n_0$, there exist functions*

$$\hat{c}_{i,j}^k \in C\big(B_{(X,d)_{J_0}}\left(x_0, \xi^1\right)\big)$$

such that, for $1 \leq i, j \leq n_0$:

$$[X_i, X_j] = \sum_{k=1}^{n_0} \hat{c}_{i,j}^k X_k.$$

These functions satisfy:

$$\sum_{|\alpha| \leq m} \left\| X_{J_0}^{\alpha} \hat{c}_{i,j}^k \right\|_{C^0\left(B_{(X,d)_{J_0}}(x_0,\xi^1)\right)} \lesssim_m 1.$$

Proof. For $1 \leq j, k \leq q$, let $X^{(j,k)}$ be the matrix obtained by replacing the jth column of the matrix X with X_k. Note that:

$$\det X_{I_0,J_0}^{(j,k)} = \epsilon_{j,k} \det X_{I_0,J(j,k)}$$

where $\epsilon_{j,k} \in \{0, 1, -1\}$, and $J(j,k) \in \mathcal{I}(n_0, q)$.

Thus, for any $1 \leq k \leq q$, we may write, by Cramer's rule:

$$X_k = \sum_{l=1}^{n_0} \frac{\det X_{I_0,J_0}^{(l,k)}}{\det X_{I_0,J_0}} X_l = \sum_{l=1}^{n_0} \epsilon_{l,k} \frac{\det X_{I_0,J(l,k)}}{\det X_{I_0,J_0}} X_l. \tag{4.4}$$

Hence, we have, for $1 \leq i, j \leq n_0$:

$$[X_i, X_j] = \sum_{k=1}^{q} c_{i,j}^k X_k = \sum_{l=1}^{n_0} \left(\sum_{k=1}^{q} c_{i,j}^k \epsilon_{l,k} \frac{\det X_{I_0,J(l,k)}}{\det X_{I_0,J_0}} \right) X_l =: \sum_{l=1}^{n_0} \hat{c}_{i,j}^l X_l.$$

Given the form of $\hat{c}_{i,j}^k$, the desired estimates on the derivatives follow immediately from Lemma 4.13. ■

We now apply Theorem 3.1 to the list of vector fields $(X, d)_{J_0}$ on the ball $B_{(X,d)_{J_0}}(x_0, \xi^1)$. We obtain ξ_1 and η_1 as in the statement of Theorem 4.1. We obtain

$$\left| \det_{n_0 \times n_0} d\Phi(t) \right| \approx \left| \det_{n_0 \times n_0} X_{J_0}(x_0) \right|.$$

But, we know from our initial assumptions that

$$\left| \det_{n_0 \times n_0} X_{J_0}(x_0) \right| \approx \left| \det_{n_0 \times n_0} X(x_0) \right|.$$

For $1 \leq j \leq n_0$, we have:

$$\|Y_j\|_{C^m\left(B_{n_0}(\eta_1)\right)} \lesssim_m 1.$$

Hence, to complete the proof of Theorem 4.1, we need to show the existence of ξ_2 and prove the estimates on Y_j for $n_0 < j \leq q$. We begin with the latter:

Proposition 4.15. $\|Y_k\|_{C^m\left(B_{n_0}(\eta_1)\right)} \lesssim_m 1$, *for* $n_0 < k \leq q$.

Proof. By (4.4), we see that we may write:

$$Y_k = \sum_{l=1}^{n_0} \epsilon_{l,k} \left(\frac{\det X_{I_0,J(l,k)}}{\det X_{I_0,J_0}} \circ \Phi \right) Y_l.$$

Since we already know the result for Y_l, $1 \leq l \leq n_0$, it suffices to show that:

$$\left\| \frac{\det X_{I_0,J(l,k)}}{\det X_{I_0,J_0}} \circ \Phi \right\|_{C^m\left(B_{n_0}(\eta_1)\right)} \lesssim_m 1.$$

By Proposition 3.18, it suffices to show that for $|\alpha| \leq m$,

$$\left\| Y_{J_0}^{\alpha} \frac{\det X_{I_0,J(l,k)}}{\det X_{I_0,J_0}} \circ \Phi \right\|_{C^0\left(B_{n_0}(\eta_1)\right)} \lesssim_m 1.$$

But,

$$Y_{J_0}^{\alpha} \frac{\det X_{I_0, J(l,k)}}{\det X_{I_0, J_0}} \circ \Phi = \left(X_{J_0}^{\alpha} \frac{\det X_{I_0, J(l,k)}}{\det X_{I_0, J_0}} \right) \circ \Phi.$$

From here, the result follows immediately from an application of Lemma 4.13. ■

We now conclude our proof of Theorem 4.1, with the following proposition:

Proposition 4.16. *There exists an admissible constant $\xi_2 > 0$ such that:*

$$B_{(X,d)}(x_0, \xi_2) \subseteq B_{(X,d)_{J_0}}(x_0, \xi_1).$$

Proof. Suppose $y \in B_{(X,d)}(x_0, \xi_2)$, where $\xi_2 \leq \xi_1 \leq \xi^1$ will be chosen at the end of the proof. Thus there exists a path $\gamma : [0,1] \to B_{(X,d)}(x_0, \xi_2)$, $\gamma(0) = x_0$, $\gamma(1) = y$,

$$\gamma'(t) = (b \cdot X)(\gamma(t))$$

where $b \in L^{\infty}([0,1])^q$ with $\left\| \left| \xi_2^{-d} b \right| \right\|_{L^{\infty}([0,1])} < 1$. Then, applying (4.4), we have:

$$\begin{aligned} \gamma'(t) &= \sum_{k=1}^{q} b_k(t) X_k(\gamma(t)) \\ &= \sum_{l=1}^{n_0} \left(\sum_{k=1}^{q} \epsilon_{l,k} b_k(t) \frac{\det X(\gamma(t))_{I_0, J(l,k)}}{\det X(\gamma(t))_{I_0, J_0}} \right) X_l(\gamma(t)) \\ &=: \sum_{l=1}^{n_0} a_l(t) X_l(\gamma(t)), \end{aligned}$$

and if $\xi_2 > 0$ is admissibly small enough, by Lemma 4.12, we have that:

$$\left\| \sqrt{\sum_{l=1}^{n_0} \xi_1^{-2d_l} |a_l|^2} \right\|_{L^{\infty}([0,1])} < 1,$$

proving that $y = \gamma(1) \in B_{(X,d)_{J_0}}(x_0, \xi_1)$. ■

4.1. Control of vector fields

We take all the same notation as in Section 4, and define (m-)admissible constants in the same way.[22] The goal of this section is to understand when we can add an additional vector field with a formal degree (X_{q+1}, d_{q+1})

[22] We are still assuming $J_0 = (1, \dots, n_0)$.

($d_{q+1} \in (0,\infty)$) to the list of vector fields (X,d) without "adding anything new." In particular, we wish to not significantly increase the size of $B_{(X,d)}(x_0,\tau)$, where τ is thought of as a fixed constant $\leq \xi$.

Let X_{q+1} be a C^1 vector field on $B_{(X,d)_{J_0}}(x_0,\xi)$ (here we mean that X_{q+1} is C^1 thought of as a function on the leaf in which $B_{(X,d)_{J_0}}(x_0,\xi)$ lies; but it need not be tangent to the leaf), and assign to it a formal degree $d_{q+1} \in (0,\infty)$. Let $\left(\widehat{X},\hat{d}\right)$ denote the list of vector fields with formal degrees:

$$((X_1,d_1),\ldots,(X_{q+1},d_{q+1})).$$

For an integer $m \geq 1$ we define three conditions which will turn out to be equivalent (all parameters below are considered to be elements of $(0,\infty)$):

1. $\mathcal{P}_1^m(\kappa_1,\tau_1,\sigma_1,\sigma_1^m)$:
   - $\left|\det_{n_0\times n_0} X(x_0)\right|_\infty \geq \kappa_1 \left|\det_{n_0\times n_0}\widehat{X}(x_0)\right|_\infty$
   - $\left|\det_{j\times j}\widehat{X}(x_0)\right| = 0$, $n_0 < j \leq n$.
   - There exist $c_{i,q+1}^j \in C^0\big(B_{(X,d)_{J_0}}(x_0,\tau_1)\big)$ such that

     $$[X_i,X_{q+1}] = \sum_{j=1}^{q+1} c_{i,q+1}^j X_j, \quad \text{on } B_{(X,d)_{J_0}}(x_0,\tau_1)$$

     with:

     $$\sum_{|\alpha|\leq m-1}\left\|X^\alpha c_{i,q+1}^j\right\|_{C^0\left(B_{(X,d)_{J_0}}(x_0,\tau_1)\right)} \leq \sigma_1^m,\ \left\|c_{i,q+1}^j\right\|_{C^0\left(B_{(X,d)_{J_0}}(x_0,\tau_1)\right)} \leq \sigma_1.$$

2. $\mathcal{P}_2^m(\tau_2,\sigma_2,\sigma_2^m)$: There exist $c_j \in C^0\big(B_{(X,d)_{J_0}}(x_0,\tau_2)\big)$ such that:
   - $X_{q+1} = \sum_{j=1}^{n_0} c_j X_j$, on $B_{(X,d)_{J_0}}(x_0,\tau_2)$.
   - $\sum_{|\alpha|\leq m}\left\|X^\alpha c_j\right\|_{C^0\left(B_{(X,d)_{J_0}}(x_0,\tau_2)\right)} \leq \sigma_2^m$.
   - $\sum_{|\alpha|\leq 1}\left\|X^\alpha c_j\right\|_{C^0\left(B_{(X,d)_{J_0}}(x_0,\tau_2)\right)} \leq \sigma_2$.

3. $\mathcal{P}_3^m(\tau_3,\sigma_3,\sigma_3^m)$: There exist $c_j \in C^0\big(B_{(X,d)_{J_0}}(x_0,\tau_3)\big)$ such that:
   - $X_{q+1} = \sum_{j=1}^{q} c_j X_j$, on $B_{(X,d)_{J_0}}(x_0,\tau_3)$.
   - $\sum_{|\alpha|\leq m}\left\|X^\alpha c_j\right\|_{C^0\left(B_{(X,d)_{J_0}}(x_0,\tau_3)\right)} \leq \sigma_3^m$.
   - $\sum_{|\alpha|\leq 1}\left\|X^\alpha c_j\right\|_{C^0\left(B_{(X,d)_{J_0}}(x_0,\tau_3)\right)} \leq \sigma_3$.

Theorem 4.17. $\mathcal{P}_1^m \Rightarrow \mathcal{P}_2^m \Rightarrow \mathcal{P}_3^m \Rightarrow \mathcal{P}_1^m$ *in the following sense:*

1. $\mathcal{P}_1^m(\kappa_1, \tau_1, \sigma_1, \sigma_1^m) \Rightarrow$ *there exist admissible constants* $\tau_2 = \tau_2(\kappa_1, \tau_1, \sigma_1)$, $\sigma_2 = \sigma_2(\kappa_1, \sigma_1)$, *and an m-admissible constant* $\sigma_2^m = \sigma_2^m(\kappa_1, \sigma_1^m)$ *such that* $\mathcal{P}_2^m(\tau_2, \sigma_2, \sigma_2^m)$.

2. $\mathcal{P}_2^m(\tau_2, \sigma_2, \sigma_2^m) \Rightarrow \mathcal{P}_3^m(\tau_2, \sigma_2, \sigma_2^m)$.

3. $\mathcal{P}_3^m(\tau_3, \sigma_3, \sigma_3^m) \Rightarrow$ *there exist admissible constants* $\kappa_1 = \kappa_1(\sigma_3)$, $\sigma_1 = \sigma_1(\sigma_3)$ *and an m-admissible constant* $\sigma_1^m = \sigma_1^m(\sigma_3^m)$, *such that* $\mathcal{P}_1^m(\kappa_1, \tau_3, \sigma_1, \sigma_1^m)$.

Proof. $\mathcal{P}_1^m \Rightarrow \mathcal{P}_2^m$ follows just as in the proof of (4.4). This can be seen by noting that (X, d) can be replaced by $\left(\widehat{X}, \hat{d}\right)$ in the proofs of Lemmas 4.10, 4.11, 4.12, and 4.13. The reader might worry that in the definition of $\mathcal{P}_2^m$ we are using X^α instead of $X_{J_0}^\alpha$; however, there is no real difference between the two, due to (4.4). From there, the proof follows easily, and we leave the details to the interested reader. $\mathcal{P}_2^m \Rightarrow \mathcal{P}_3^m$ and $\mathcal{P}_3^m \Rightarrow \mathcal{P}_1^m$ are both trivial. ∎

Let $d_{q+1}^\vee$ be a fixed lower bound for d_{q+1}. We have:

Proposition 4.18. *Suppose* $\mathcal{P}_2^1(\tau_2, \sigma_2, \sigma_2^1)$ *holds. Then, there exists an admissible constant* $\tau' = \tau'\left(d_{q+1}^\vee, \tau_2, \sigma_2\right)$ *such that:*

$$B_{(X,d)}(x_0, \tau') \subseteq B_{\left(\widehat{X}, \hat{d}\right)}(x_0, \tau') \subseteq B_{(X,d)_{J_0}}(x_0, \tau_2).$$

Proof. The first containment is trivial. The second follows just as in the proof of Proposition 4.16. ∎

Proposition 4.19. *Suppose* $\mathcal{P}_2^m(\tau_2, \sigma_2, \sigma_2^m)$ *holds. Let* $\eta' \leq \eta_1$ *be small enough that* $\Phi(B_{n_0}(\eta')) \subseteq B_{(X,d)_{J_0}}(x_0, \tau_2)$. *Let* Y_{q+1} *be the pullback of* X_{q+1} *under* Φ *to* $B_{n_0}(\eta')$. *Then,*

$$\|Y_{q+1}\|_{C^m\left(B_{n_0}(\eta')\right)} \leq \sigma_4^m$$

where $\sigma_4^m = \sigma_4^m(\sigma_2^m)$ *is an m-admissible constant.*

Proof. This follows just as in Proposition 4.15. ∎

Remark 4.20. Our assumption on the commutator $[X_i, X_j]$ in Section 4 was essentially just that $([X_i, X_j], d_i + d_j)$ satisfied condition $\mathcal{P}_3^m$ for appropriate m.

4.2. Unit operators at the unit scale

In this section, we study the compositions of certain "unit operators," which will be the core of our study of maximal functions in Section 6–see Section 1.2.4 for some motivation for the study of these operators.

Let $X_1, \ldots, X_q$, $d_1, \ldots, d_q$, x_0, ξ, ζ, $c_{i,j}^k$, n_0, and J_0 be as in Section 4, in addition (for simplicity), we assume the stronger assumptions of Remark 4.5. We again suppose, without loss of generality, that $J_0 = (1, \ldots, n_0)$. In addition, suppose we are given ν subsets of $\{(X_1, d_1), \ldots, (X_q, d_q)\}$:

$$\left\{(Z_1^\mu, d_1^\mu), \ldots, \left(Z_{q_\mu}^\mu, d_{q_\mu}^\mu\right)\right\} \subseteq \{(X_1, d_1), \ldots, (X_q, d_q)\}$$

with $1 \leq \mu \leq \nu$. Suppose these subsets satisfy:

$$\{(X_1, d_1), \ldots, (X_{n_0}, d_{n_0})\} \subseteq \bigcup_{1\leq\mu\leq\nu} \left\{(Z_1^\mu, d_1^\mu), \ldots, \left(Z_{q_\mu}^\mu, d_{q_\mu}^\mu\right)\right\}. \tag{4.5}$$

We say C is a pre-admissible constant if C can be chosen to depend only on those parameters an admissible constant could depend on in Remark 4.5, plus a fixed upper bound for ν. We will write $A \lesssim B$ for $A \leq CB$ where C is a pre-admissible constant. Also, we write $A \simeq B$ for $A \lesssim B$ and $B \lesssim A$.

If we say that C is an admissible constant, it means that we furthermore assume that:

$$\left[Z_i^\mu, Z_j^\mu\right] = \sum_{k=1}^{q_\mu} c_{i,j}^{k,\mu} Z_k^\mu$$

and C is allowed to depend on everything a pre-admissible constant is allowed to depend on, plus a fixed upper bound for the quantities:

$$\sum_{|\alpha|\leq 2} \left\|(Z^\mu)^\alpha c_{i,j}^{k,\mu}\right\|_{C^0\left(B_{(X,d)}(X_0,\xi)\right)}, \quad 1 \leq \mu \leq \nu$$

which we assume to exist and are finite.

Given a function f defined on a set U, and given for each $x \in U$ a set V_x, we define for those $y \in U$ such that $V_y \subseteq U$:

$$A_{U,V_\cdot} f(y) = \frac{1}{\mathrm{Vol}(V_y)} \int_{V_y} f(z)\, dz.$$

Here we are being ambiguous about what we mean by $\mathrm{Vol}(V_y)$ and dz. Below, V will be replaced by sets lying in the leaf generated by one of the Z^μs (or by X). We then mean for $\mathrm{Vol}(\cdot)$ and dz to refer to the Lebesgue measure on that leaf. Below, we will drop the U from the subscript $A_{U,V_\cdot}$, and it is understood to be the domain of f.

If we let ξ_2 be as in the statement of Corollary 4.2[23] the main result of this section is:

Theorem 4.21. *There exist admissible constants* $0 < \lambda_3, \lambda_2, \lambda_1 \leq \xi_2$ *such that for every* $f \in C^0\left(B_{(X,d)}(x_0, \xi)\right)$ *with* $f \geq 0$*, we have:*

$$A_{B_{(X,d)}(\cdot,\lambda_3)} f(x_0) \lesssim A_{B_{(Z^\nu,d^\nu)}(\cdot,\lambda_2)} A_{B_{(Z^{\nu-1},d^{\nu-1})}(\cdot,\lambda_2)} \cdots A_{B_{(Z^1,d^1)}(\cdot,\lambda_2)} f(x_0)$$
$$\lesssim A_{B_{(X,d)}(\cdot,\lambda_1)} f(x_0).$$

Define $n_1 = \sum_{\mu=1}^{\nu} q_\mu$. To prove Theorem 4.21, we need a preliminary result:

Proposition 4.22. *There exist pre-admissible constants* $0 < l_3, l_2, l_1 \leq \xi_2$ *such that for all* $f \geq 0$*, we have:*

$$A_{B_{(X,d)}(\cdot,l_3)} f(x_0) \precsim \int_{Q_{n_1}(l_2)} f\left(e^{u_1\cdot Z^1} e^{u_2\cdot Z^2} \cdots e^{u_\nu \cdot Z^\nu} x_0\right) du_1 \dots du_\nu du_\nu$$
$$\precsim A_{B_{(X,d)}(\cdot,l_1)} f(x_0).$$

Recall, $Q_{n_1}(l_2)$ *denotes the* $|\cdot|_\infty$ *ball in* $\mathbb{R}^{n_1}$ *of radius* l_2.

To prove Proposition 4.22, we need some preliminary results. Set $l_1 = \xi_2$, take η_2 and Φ as in Corollary 4.2.

Lemma 4.23. *For* $f \geq 0$,

$$\int_{B_{n_0}(\eta_2)} f \circ \Phi(u)\, du \precsim A_{B_{(X,d)}(\cdot,l_1)} f(x_0).$$

Proof. Note that:

$$\mathrm{Vol}\left(\Phi\left(B_{n_0}(\eta_2)\right)\right) \simeq \left|\det_{n_0\times n_0} X(x_0)\right| \simeq \mathrm{Vol}\left(B_{(X,d)}(x_0, \lambda_1)\right)$$

and so we have:

$$\frac{1}{\mathrm{Vol}\left(\Phi\left(B_{n_0}(\eta_2)\right)\right)} \int_{\Phi\left(B_{n_0}(\eta_2)\right)} f(y)\, dy \precsim A_{B_{(X,d)}(\cdot,l_1)} f(x_0).$$

Now applying a change of variables as in (B.2) and using that by Theorem 4.1 for every $u \in B_{n_0}(\eta_2)$,

$$\left|\det_{n_0\times n_0} d\Phi(u)\right| \simeq \left|\det_{n_0\times n_0} X(x_0)\right| \simeq \mathrm{Vol}\left(B_{(X,d)}(x_0, l_1)\right)$$

we see that:

$$\int_{B_{n_0}(\eta_2)} f \circ \Phi(u)\, du \precsim A_{B_{(X,d)}(\cdot,l_1)} f(x_0)$$

completing the proof. ∎

[23]Note that all of the constants in Corollary 4.2 are pre-admissible in the sense of this section.

Let Y_j^μ be the pullback of Z_j^μ under the map Φ. As before, we let $Y_1, \ldots, Y_{n_0}$ denote the pullbacks of $X_1, \ldots, X_{n_0}$. Note that, by (4.5) each of $Y_1, \ldots, Y_{n_0}$ appears as at least one of the Y_j^μ.

Lemma 4.24. *There exists pre-admissible constants $l_2 > 0$, $\eta_3 > 0$ such that for all $f \in C^0(B_{n_0}(\eta_2))$, $f \geq 0$,*

$$\int_{B_{n_0}(\eta_3)} f(u)\, du \lesssim \int_{Q_{n_1}(l_2)} f\left(e^{u_1\cdot Y^1} e^{u_2\cdot Y^2} \cdots e^{u_\nu \cdot Y^\nu} 0\right) du_1 \cdots du_\nu \lesssim \int_{B_{n_0}(\eta_2)} f(u)\, du. \tag{4.6}$$

Proof. Let $\Psi(u_1, \ldots u_\nu)$ denote the map:

$$\Psi(u_1, \ldots, u_\nu) = e^{u_1\cdot Y^1} e^{u_2\cdot Y^2} \cdots e^{u_\nu \cdot Y^\nu} 0.$$

Note that $\Psi \in C^2(Q_{n_1}(\eta'))$, provided η' is a sufficiently small pre-admissible constant. Moreover, the C^2 norm is bounded by a pre-admissible constant (by Theorem A.1, using that $\left\|Y_j^\mu\right\|_{C^2\left(B_{n_0}(\eta_2)\right)} \lesssim 1$, by Theorem 4.1).

Recalling that each Y_j $(1 \leq j \leq n_0)$ appears at least once in some Y_k^μ, for each $1 \leq j \leq n_0$ we pick one such occurrence. Write Ψ as a function of two variables:

$$\Psi\left(u^1, u^2\right), \quad u^1 \in Q_{n_0}(\eta'), u^2 \in Q_{n_1 - n_0}(\eta')$$

where u^1 denotes the coefficients of the above chosen Y_j, and u^2 denotes the remaining coefficients. For each fixed u^2, think of Ψ as a function of one variable:

$$\Psi_{u^2}\left(u^1\right)$$

Note $d\Psi_0(0) = I$, and so by the C^2 estimates of Ψ, we see that if l_2 if a pre-admissible constant that is small enough, for every $u^2 \in Q_{n_1-n_0}(l_2)$, we have:

$$\left\|d\Psi_{u^2}(0) - I\right\| \leq \frac{1}{2}. \tag{4.7}$$

Hence, by the inverse function theorem (Theorem A.3), we may pre-admissibly shrink l_2 such that:

- For every $u^2 \in Q_{n_1-n_0}(l_2)$, Ψ_{u^2} is injective on $Q_{n_0}(l_2)$.
- $\Psi_{u^2}(Q_{n_0}(l_2)) \subseteq Q_{n_0}(\eta_2)$, for every $u^2 \in Q_{n_1-n_0}(l_2)$.

Hence, by a simple change of variables, we have for $u^2 \in Q_{n_1-n_0}(l_2)$:

$$\int_{Q_{n_0}(l_2)} f\left(\Psi_{u^2}\left(u^1\right)\right) du^1 \lesssim \int_{B_{n_0}(\eta_2)} f(u)\, du$$

for $f \geq 0$. Applying:

$$\int_{Q_{n_1-n_0}(l_2)} du^2$$

to both sides of this expression proves the latter inequality in (4.6).

We now turn to the former inequality in (4.6). Applying the inverse function theorem (Theorem A.3) and again using (4.7) we have that there exist pre-admissible constants $\eta' \leq l_2$ and $\eta_3 \leq l_2$ such that for every $u^2 \in Q_{n_1-n_0}(\eta')$ we have that:

$$B_{n_0}(\eta_3) \subseteq \Psi_{u^2}(Q_{n_0}(l_2)).$$

Thus a simple change of variables (using (4.7)) shows that, for $u^2 \in Q_{n_1-n_0}(\eta')$ and $f \geq 0$:

$$\int_{B_{n_0}(\eta_3)} f(u)\, du \lesssim \int_{Q_{n_0}(l_2)} f\left(\Psi_{u^2}\left(u^1\right)\right) du^1.$$

Integrating both sides in u^2, we obtain:

$$\begin{aligned}\int_{B_{n_0}(\eta_3)} f(u)\, du &\lesssim \int_{Q_{n_1-n_0}(\eta')} \int_{Q_{n_0}(l_2)} f\left(\Psi_{u^2}\left(u^1\right)\right) du^1 du^2 \\ &\lesssim \int_{Q_{n_1-n_0}(l_2)} \int_{Q_{n_0}(l_2)} f\left(\Psi_{u^2}\left(u^1\right)\right) du^1 du^2.\end{aligned}$$

Where in the last line, we used that $f \geq 0$ and that $\eta' \leq l_2$. This proves the first inequality in (4.6) and completes the proof. ■

Lemma 4.25. *There exists a pre-admissible constant $l_3 > 0$ such that for all $f \geq 0$:*

$$A_{B_{(X,d)}(\cdot,l_3)} f(x_0) \lesssim \int_{B_{n_0}(\eta_3)} f \circ \Phi(u)\, du.$$

Proof. Proceeding as in the proofs of Propositions 3.21 and 4.16, we may find a pre-admissible constant $l_3 > 0$ such that:

$$B_{(X,d)}(x_0, l_3) \subseteq \Phi(B_{n_0}(\eta_3)).$$

Note that, by Remark 4.9, we have:

$$\mathrm{Vol}\left(B_{(X,d)}(x_0, l_3)\right) \simeq \left| \det_{n_0 \times n_0} X(x_0) \right| \simeq \mathrm{Vol}\left(\Phi\left(B_{n_0}(\eta_3)\right)\right)$$

and it follows that

$$A_{B_{(X,d)}(\cdot,l_3)}f(x_0) \lesssim \frac{1}{\operatorname{Vol}(\Phi(B_{n_0}(\eta_3)))} \int_{\Phi\left(B_{n_0}(\eta_3)\right)} f(y)\,dy.$$

Applying a change of variables as in (B.2) and using that for all $u \in B_{n_0}(\eta_3)$, we have:

$$\left|\det_{n_0\times n_0} d\Phi(t)\right| \simeq \left|\det_{n_0\times n_0} X(x_0)\right| \simeq \operatorname{Vol}(\Phi(B_{n_0}(\eta_3)))$$

it follows that:

$$A_{B_{(X,d)}(\cdot,l_3)}f(x_0) \lesssim \int_{B_{n_0}(\eta_3)} f(u)\,du$$

completing the proof. ■

Proof of Proposition 4.22. Fix $f \geq 0$ as in the statement of Proposition 4.22. Apply Lemmas 4.23, 4.25 to f and Lemma 4.24 to $f \circ \Phi$ to obtain:

$$\begin{aligned} A_{B_{(X,d)}(\cdot,l_3)}f(x_0) &\lesssim \int_{Q_{n_1}(l_2)} f\circ\Phi\left(e^{u_1\cdot Y^1}e^{u_2\cdot Y^2}\cdots e^{u_\nu\cdot Y^\nu}0\right) du_1\ldots du_\nu du_\nu \\ &\lesssim A_{B_{(X,d)}(\cdot,l_1)}f(x_0). \end{aligned}$$

Using that:

$$f\circ\Phi\left(e^{u_1\cdot Y^1}e^{u_2\cdot Y^2}\cdots e^{u_\nu\cdot Y^\nu}0\right) = f\left(e^{u_1\cdot Z^1}e^{u_2\cdot Z^2}\cdots e^{u_\nu\cdot Z^\nu}x_0\right)$$

completes the proof. ■

Proof of Theorem 4.21. Let $0 < \xi' \leq \xi_2$ be an admissible constant so small that:

$$\begin{aligned} \Omega_0 := &\bigcup_{\substack{x_\nu\in\\ B_{(Z^\nu,d^\nu)}(x_0,\xi')}} \bigcup_{\substack{x_{\nu-1}\in\\ B_{(Z^{\nu-1},d^{\nu-1})}(x_\nu,\xi')}} \cdots \bigcup_{\substack{x_2\in\\ B_{(Z^2,d^2)}(x_3,\xi')}} B_{(Z^1,d^1)}(x_2,\xi') \\ &\Subset B_{(X,d)}\left(x_0,\frac{\xi}{2}\right), \end{aligned}$$

where $A \Subset B$ denotes that A is a relatively compact subset of B. It is easy to see that this is possible, and we leave the details to the reader. Further, we take $0 < \xi'' \leq \xi'$ to be an admissible constant so small that for every $y \in \Omega_0$,

$$B_{(Z^\mu,d^\mu)}(y,\xi'') \Subset B_{(X,d)}(x_0,\xi), \quad 1 \leq \mu \leq \nu.$$

We apply Proposition 4.22 to each $y \in \Omega_0$ with ξ'' in place of ξ and (Z^μ, d^μ) in place of (X, d) (and taking $\nu = 1$) to find admissible constants l_1, l_2, l_3 such that for every $y \in \Omega_0$, and every $f \geq 0$

$$A_{B_{(Z^\mu,d^\mu)}(\cdot,l_3)} f(y) \lesssim \int_{Q_{q_\mu}(l_2)} f\left(e^{u_\mu \cdot Z^\mu} y\right) du_\mu \lesssim A_{B_{(Z^\mu,d^\mu)}(\cdot,l_1)} f(y) \tag{4.8}$$

and also applying Proposition 4.22 as it is stated we may ensure that:

$$\begin{aligned} A_{B_{(X,d)}(\cdot,l_3)} f(x_0) &\lesssim \int_{Q_{n_1}(l_2)} f\left(e^{u_1 \cdot Z^1} e^{u_2 \cdot Z^2} \cdots e^{u_\nu \cdot Z^\nu} x_0\right) du_1 \ldots du_\nu du_\nu \\ &\lesssim A_{B_{(X,d)}(\cdot,l_1)} f(x_0). \end{aligned} \tag{4.9}$$

Let $\lambda_1 = l_1$ and $\lambda_2 = l_3$. Then, applying (4.8) ν times, we see that:

$$\begin{aligned} &A_{B_{(Z^\nu,d^\nu)}(\cdot,\lambda_2)} A_{B_{(Z^{\nu-1},d^{\nu-1})}(\cdot,\lambda_2)} \cdots A_{B_{(Z^1,d^1)}(\cdot,\lambda_2)} f(x_0) \lesssim \\ &\lesssim \int_{Q_{n_1}(l_2)} f\left(e^{u_1 \cdot Z^1} e^{u_2 \cdot Z^2} \cdots e^{u_\nu \cdot Z^\nu} x_0\right) du_1 \ldots du_\nu. \end{aligned}$$

Applying (4.9) yields the second inequality in the statement of Theorem 4.21.

We apply Proposition 4.22 to each $y \in \Omega_0$ with λ_2 in place of ξ and (Z^μ, d^μ) in place of (X, d) (and taking $\nu = 1$) to find admissible constants $l_3', l_2', l_1' \leq \lambda_2$ such that for every $y \in \Omega_0$, and every $f \geq 0$:

$$A_{B_{(Z^\mu,d^\mu)}\left(\cdot,l_3'\right)} f(y) \lesssim \int_{Q_{q_\mu}\left(l_2'\right)} f\left(e^{u_\mu \cdot Z^\mu} y\right) du_\mu \lesssim A_{B_{(Z^\mu,d^\mu)}\left(\cdot,l_1'\right)} f(y) \tag{4.10}$$

and also applying Proposition 4.22 as it is stated (with λ_2 in place of ξ) we may ensure that:

$$\begin{aligned} A_{B_{(X,d)}\left(\cdot,l_3'\right)} f(x_0) &\lesssim \int_{Q_{n_1}\left(l_2'\right)} f\left(e^{u_1 \cdot Z^1} e^{u_2 \cdot Z^2} \cdots e^{u_\nu \cdot Z^\nu} x_0\right) du_1 \ldots du_\nu du_\nu \\ &\lesssim A_{B_{(X,d)}\left(\cdot,l_1'\right)} f(x_0). \end{aligned} \tag{4.11}$$

Set $\lambda_3 = l_3'$. We first claim that, for all $f \geq 0$:

$$A_{B_{(Z^\mu,d^\mu)}\left(\cdot,l_1'\right)} f(y) \lesssim A_{B_{(Z^\mu,d^\mu)}(\cdot,\lambda_2)} f(y), \quad y \in \Omega_0, \quad 1 \leq \mu \leq \nu. \tag{4.12}$$

Indeed, we already have that $l_1' \leq \lambda_2$. Moreover, we have by Remark 4.9:

$$\mathrm{Vol}\left(B_{(Z^\mu,d^\mu)}\left(y, l_1'\right)\right) \approx \mathrm{Vol}\left(B_{(Z^\mu,d^\mu)}\left(y, \lambda_2\right)\right), \quad y \in \Omega_0$$

and (4.12) immediately follows.

Thus we have:

$$
\begin{aligned}
A_{B_{(Z^\nu,d^\nu)}(\cdot,\lambda_2)} A_{B_{(Z^{\nu-1},d^{\nu-1})}(\cdot,\lambda_2)} \cdots A_{B_{(Z^1,d^1)}(\cdot,\lambda_2)} f(x_0) &\gtrsim \\
&\gtrsim A_{B_{(Z^\nu,d^\nu)}(\cdot,l_1')} A_{B_{(Z^{\nu-1},d^{\nu-1})}(\cdot,l_1')} \cdots A_{B_{(Z^1,d^1)}(\cdot,l_1')} f(x_0) \\
&\gtrsim \int_{Q_{n_1}(l_2')} f\left(e^{u_1\cdot Z^1} e^{u_2\cdot Z^2} \cdots e^{u_\nu\cdot Z^\nu} x_0\right) du_1 \ldots du_\nu du_\nu \\
&\gtrsim A_{B_{(X,d)}(\cdot,\lambda_3)} f(x_0)
\end{aligned}
$$

where in the second to last line, we have applied (4.10) ν times, and in the last line we have applied (4.11). This completes the proof. ■

5. Multi-parameter Carnot-Carathéodory balls

In this section, we discuss multi-parameter Carnot-Carathéodory balls. In Section 5.1 we state the main theorem regarding multi-parameter balls (Theorem 5.3). In Section 5.2 we discuss four examples/applications where Theorem 5.3 applies, one of which is the "weakly-comparable" balls of [30]. Finally, in Section 5.3 we discuss a notion of "controlling" vector fields, which we hope will elucidate the complicated assumptions in Section 5.1.

Before we begin, we need one new piece of notation. Suppose we are given formal degrees $d_1, \ldots, d_q \in [0,\infty)^\nu$. If α is an ordered multi-index, we define the formal degree

$$d(\alpha) = \sum_{j=1}^{q} k_j d_j$$

where k_j denotes the number of times that j appears in the list α. Thus if $\delta \in [0,\infty)^\nu$, we may define $\delta^{d(\alpha)} \in [0,\infty)$ and $\delta^{-d(\alpha)} \in [0,\infty]$ in the usual way.

5.1. The main theorem

Suppose $X_1, \ldots, X_q$ are q C^1 vector fields with associated formal degrees $0 \neq d_1, \ldots, d_q \in [0,\infty)^\nu$. Let $K \subset \Omega$ (think of $K = \{x_0\}$ or, more generally, K compact). Suppose that $\xi \in (0,1]^\nu$ is such that (X,d) satisfies $\mathcal{C}(x,\xi)$, for every $x \in K$. The goal in this section is to apply Theorem 4.1 and Corollaries 4.2 and 4.3 to the vector fields $(\delta X, \sum d)$ at each point $x \in K$, where $\delta \in [0,1)^\nu$ is small.

Fix a subset $\mathcal{A}$:

$$\mathcal{A} \subseteq \{\delta \in [0,1]^\nu : \delta \neq 0, \delta \leq \xi\}$$

to be the set of "allowable" δs. Recall, $\delta \leq \xi$ means that the inequality holds coordinatewise.

Remark 5.1. We will be restricting our attention to balls $B_{(X,d)}(x,\delta)$, where $\delta \in \mathcal{A}$, $x \in K$. For many applications, one would take:

$$\mathcal{A} = \{\delta \in [0,1]^\nu : \delta \neq 0, \delta \leq \xi\} \tag{5.1}$$

and we encourage the reader to keep this particular choice of $\mathcal{A}$ in mind throughout this section. However, other choices of $\mathcal{A}$ do arise in applications. For instance, the choice:

$$\mathcal{A} = \{\delta \in [0,1]^\nu : \delta \neq 0, \delta \leq \xi, \delta_1 \geq \delta_2 \geq \cdots \geq \delta_\nu\} \tag{5.2}$$

arises in the study of flag kernels, as in [17]. Also, the results in Section 5.2.1 use yet another choice of $\mathcal{A}$.

In this section, we assume that for every $\delta \in \mathcal{A}$, $x \in K$, we have:

$$\left[\delta^{d_i} X_i, \delta^{d_j} X_j\right] = \sum_k c_{i,j}^{k,\delta,x} \delta^{d_k} X_k$$

on $B_{(X,d)}(x,\delta)$. In addition, we assume:

- The X_js are C^2 on $B_{(X,d)}(x,\xi)$, for every $x \in K$, and satisfy
$$\sup_{x\in K} \|X_j\|_{C^2\left(B_{(X,d)}(x,\xi)\right)} < \infty\,.$$
- For all $|\alpha| \leq 2$, $x \in K$, we have $\left(\delta^d X\right)^\alpha c_{i,j}^{k,\delta,x} \in C^0\left(B_{(X,d)}(x,\delta)\right)$, for every i,j,k, and every $\delta \in \mathcal{A}$, and moreover:
$$\sup_{\substack{\delta\in A\\ x\in K}} \sum_{|\alpha|\leq 2} \left\|\left(\delta^d X\right)^\alpha c_{i,j}^{k,\delta,x}\right\|_{C^0\left(B_{(X,d)}(x,\delta)\right)} < \infty.$$

Finally, let

$$n_0(x,\delta) = \dim \operatorname{span}\left\{\delta^{d_1} X_1(x), \ldots, \delta^{d_q} X_q(x)\right\}.$$

We say C is an admissible constant if C can be chosen to depend only on fixed upper and lower bounds $d_{max} < \infty$, $d_{min} > 0$, for the coordinates of $\sum d$, a fixed upper bound for n, q, ν and a fixed upper bound for the quantities:

$$\sup_{x\in K} \|X_j\|_{C^2\left(B_{(X,d)}(x,\xi)\right)}, \quad \sup_{\substack{\delta\in A\\ x\in K}} \sum_{|\alpha|\leq 2} \left\|\left(\delta^d X\right)^\alpha c_{i,j}^{k,\delta,x}\right\|_{C^0\left(B_{(X,d)}(x,\delta)\right)}.$$

Furthermore, if we say C is an m-admissible constant, we mean that in addition to the above, we assume that:

- $\sup_{x\in K}\|X_j\|_{C^m\left(B_{(X,d)}(x,\xi)\right)}<\infty$, for every $1\leq j\leq q$.
- $\sup_{\substack{\delta\in A\\ x\in K}}\sum_{|\alpha|\leq m}\left\|\left(\delta^d X\right)^\alpha c_{i,j}^{k,\delta,x}\right\|_{C^0\left(B_{(X,d)}(x,\delta)\right)}<\infty$, for every i,j,k.

(in particular, the above partial derivatives exist and are continuous). C is allowed to depend on m, all the quantities an admissible constant is allowed to depend on, and a fixed upper bound for the above two quantities.

Remark 5.2. The assumptions in this section are somewhat complicated. The reader might hope that special cases of these assumptions might be enough for applications. Unfortunately, this seems to not be the case, and is discussed in Section 5.3.

For each $\delta\in\mathcal{A}$, $x\in K$, let $J(x,\delta)=\left(J(x,\delta)_1,\ldots,J(x,\delta)_{n_0(x,\delta)}\right)\in\mathcal{I}(n_0(x,\delta),q)$ be such that:

$$\left|\det_{n_0(x,\delta)\times n_0(x,\delta)}\left(\delta^d X(x)\right)_{J(x,\delta)}\right|_\infty=\left|\det_{n_0(x,\delta)\times n_0(x,\delta)}\delta^d X(x)\right|_\infty,$$

and define, for $u\in\mathbb{R}^{n_0(x,\delta)}$ with $|u|$ sufficiently small:

$$\Phi_{x,\delta}(u)=e^{u\cdot\left(\delta^d X\right)_{J(x,\delta)}}x.$$

The main result of this section is:

Theorem 5.3. *There exist admissible constants $\eta_1,\eta_2>0$, $0<\xi_4\leq\xi_3<\xi_2\leq\xi_1$ such that, for all $\delta\in\mathcal{A}$, $x\in K$:*

$$\begin{aligned}B_{(X,d)}(x,\xi_4\delta)&\subseteq B_{(X,d)_{J(x,\delta)}}(x,\xi_3\delta)\subseteq\Phi_{x,\delta}\left(B_{n_0(x,\delta)}(\eta_2)\right)\\&\subseteq\widetilde{B}_{(X,d)_{J(x,\delta)}}(x,\xi_2\delta)\subseteq B_{(X,d)_{J(x,\delta)}}(x,\xi_2\delta)\subseteq B_{(X,d)}(x,\xi_2\delta)\\&\subseteq B_{(X,d)_{J(x,\delta)}}(x,\xi_1\delta)\subseteq\Phi_{x,\delta}\left(B_{n_0(x,\delta)}(\eta_1)\right)\subseteq\widetilde{B}_{(X,d)_{J(x,\delta)}}(x,\delta)\\&\subseteq B_{(X,d)_{J(x,\delta)}}(x,\delta)\subseteq B_{(X,d)}(x,\delta),\end{aligned}$$

and

- $\Phi_{x,\delta}:B_{n_0(x,\delta)}(\eta_1)\to\widetilde{B}_{(X,d)_{J(x,\delta)}}(x,\delta)$ *is one-to-one.*
- *For all $u\in B_{n_0(x,\delta)}(\eta_1)$,*

$$\left|\det_{n_0(x,\delta)\times n_0(x,\delta)}d\Phi_{x,\delta}(u)\right|\approx\left|\det_{n_0(x,\delta)\times n_0(x,\delta)}\delta^d X(x)\right|.$$

- $\mathrm{Vol}\left(B_{(X,d)}(x,\xi_2\delta)\right)\approx\left|\det_{n_0(x,\delta)\times n_0(x,\delta)}\delta^d X(x)\right|$.

- *There exists* $\phi_{x,\delta} \in C_0^2\left(B_{(X,d)}(x,\delta)\right)$, *which equals* 1 *on* $B_{(X,d)}(x,\xi_4\delta)$ *and satisfies:*
$$|X^\alpha \phi_{x,\delta}| \lesssim_{(|\alpha|-1)\vee 0} \delta^{-d(\alpha)}.$$

Furthermore, if we let $Y_j^{x,\delta}$ *be the pullback of* $\delta^{d_j} X_j$ *under the map* $\Phi_{x,\delta}$ *to* $B_{n_0(x,\delta)}(\eta_1)$, *we have that:*

$$\left\| Y_j^{x,\delta} \right\|_{C^m\left(B_{n_0(x)}(\eta_1)\right)} \lesssim_m 1.$$

Finally, if for each $x \in K$, $u \in B_{n_0(x,\delta)}(\eta_1)$, *and* $\delta \in \mathcal{A}$, *we define the* $n_0(x,\delta) \times n_0(x,\delta)$ *matrix* $A(x,u)$ *by:*

$$\left(Y^{x,\delta}_{J(x,\delta)_1}, \ldots, Y^{x,\delta}_{J(x,\delta)_{n_0(x,\delta)}}\right) = (I + A(x,\cdot)) \bigtriangledown_u$$

then,

$$\sup_{u \in B_{n_0(x,\delta)}(\eta_1)} \|A(x,u)\| \leq \frac{1}{2}. \tag{5.3}$$

Proof. For each $x \in K$ and $\delta \in \mathcal{A}$, merely apply Theorem 4.1 and Corollaries 4.2 and 4.3 to $\left(\delta^d X, \sum d\right)$, taking $\zeta = 1$ and $J_0 = J(x,\delta)$. It is easy to see, by the assumptions in this section, that all of the constants admissible (respectively, m-admissible) in those results are admissible (respectively, m-admissible) in the sense of this section. ■

Corollary 5.4. *We assume, in addition to the other assumptions in this section, that for every* $\delta \in \mathcal{A}$ *with* $|\delta|$ *sufficiently small,* $\xi_2^{-1}\delta \in \mathcal{A}$ *(in particular, this is true if* $\mathcal{A}$ *is given by* (5.1) *or* (5.2)*). We have, for* $x \in K$, *and all* $\delta \in \mathcal{A}$ *with* $|\delta|$ *sufficiently small:*

$$\mathrm{Vol}\left(B_{(X,d)}(x,\delta)\right) \approx \left| \det_{n_0(x)\times n_0(x)} \left(\xi_2^{-1}\delta\right)^d X(x) \right| \approx \left| \det_{n_0(x)\times n_0(x)} \delta^d X(x) \right| \tag{5.4}$$

and so if $|\delta|$ *is sufficiently small and* $2\delta \in \mathcal{A}$,

$$\mathrm{Vol}\left(B_{(X,d)}(x,2\delta)\right) \lesssim \mathrm{Vol}\left(B_{(X,d)}(x,\delta)\right). \tag{5.5}$$

Proof. (5.4) follows by replacing δ with $\xi_2^{-1}\delta$ in the statement of Theorem 5.3. (5.5) follows since the RHS of (5.4) is the square root of a polynomial in δ (with positive coefficients). ■

5.2. Applications and examples

In this section, we present four applications/examples where Theorem 5.3 applies. The first two applications were both previously well understood, and in fact both can be understood by the methods of [22]. The reason we include them here is to put them in the context of Theorem 5.3, and because they have been useful in the past. The third example is included to provide a simple situation where the methods of [22] do not apply but Theorem 5.3 does. We close this section with most interesting of our applications. In this application, we show how to lift results from the single parameter case to the multi-parameter case. In particular, we will see how results like the Campbell-Hausdorff formula can be applied even in the multi-parameter case–where, at first glance, they seem totally inapplicable.

5.2.1. Weakly comparable balls

In this section, we discuss the so-called "weakly-comparable" balls that were used in [30]. We do not attempt to proceed in the greatest possible generality, and instead just try to present the main ideas. Most of the conclusions of this section are contained in [30], and the main purpose here is just to show how these results are a special case of Theorem 5.3.

Let $X_1, \ldots, X_\nu$ be ν C^∞ vector fields defined on Ω, with associated formal degrees $d_1, \ldots, d_\nu \in (0, \infty)$. Fix large constants κ, N. Essentially, we will be considering the balls generated by the vector fields $\delta_\mu^{d_\mu} X_\mu$, where we restrict our attention to those $\delta = (\delta_1, \ldots, \delta_\nu)$ such that:

$$\delta_{\mu_2}^N \leq \kappa \delta_{\mu_1} \tag{5.6}$$

for every μ_1, μ_2. We call a δ satisfying (5.6) a "weakly comparable" δ.

We assume that $X_1, \ldots, X_\nu$ satisfy Hörmander's condition. That is, $X_1, \ldots, X_\nu$, along with their commutators of all orders, span that tangent space at every point of Ω. Fix $K \Subset \Omega$, a compact subset of Ω, and let $\Omega_0 \Subset \Omega$ be such that $K \Subset \Omega_0$.

Let $\hat{d}_\mu \in [0, \infty)^\nu$ be the vector that is d_μ in the μth component, and 0 in the other components. For a list (or a "word") $w = (w_1, \ldots, w_r)$ of integers $1, \ldots, \nu$ we define:

$$\hat{d}(w) = \sum_{j=1}^r \hat{d}_{w_j},$$
$$X_w = \mathrm{ad}(X_{w_1}) \,\mathrm{ad}(X_{w_2}) \cdots \mathrm{ad}\left(X_{w_{r-1}}\right) X_{w_r}.$$

As before, for a ν vector $e = (e_1, \ldots, e_\nu) \in [0, \infty)^\nu$, define $\delta^e = \prod_{\mu=1}^\nu \delta_\mu^{e_\mu}$.

By the assumption that $X_1, \ldots, X_\nu$ satisfy Hörmander's condition, and by the relative compactness of Ω_0, there exist l lists $w^1, \ldots, w^l$ such that, for every $x \in \Omega_0$:

$$T_x\Omega = \operatorname{span}\{X_{w^1}(x), \ldots, X_{w^l}(x)\}.$$

Let $d_0 = \sup_{1\leq m\leq l} \left|\hat{d}(w^m)\right|_1$. Recall, $|v|_1 = \sum_j |v_j|$. Let (X, d) denote the finite list of vector fields along with associated formal degrees given by $\left(X_w, \hat{d}(w)\right)$, where w ranges over all lists satisfying $\left|\hat{d}(w)\right|_1 \leq Nd_0$.

Take $\xi \in (0, 1]^\nu$ so small that (X, d) satisfies $\mathcal{C}(x, \xi)$ for every $x \in K$, with Ω_0 taking the place of Ω in the definition of $\mathcal{C}(x, \xi)$.

In this section, we say that C is an admissible constant if C can be chosen to depend only on a fixed upper bound for n, a fixed upper bound for ν, a fixed upper bound for N and κ, fixed upper and lower bounds for d_μ $(1 \leq \mu \leq \nu)$, a fixed upper bound for d_0, a fixed lower bound for:

$$\inf_{x\in\Omega_0} \left|\det_{n\times n}\left(X_{w^1}(x)\,|\cdots|X_{w^l}(x)\right)\right|,$$

and fixed upper bounds for a finite number of the norms:

$$\|X_\mu\|_{C^m(\Omega_0)}, \quad 1 \leq \mu \leq \nu.$$

Theorem 5.5. *Let* $\mathcal{A} = \left\{\delta \in [0, 1]^\nu : \delta \neq 0, \delta \leq \xi, \delta_{\mu_2}^N \leq \kappa\delta_{\mu_1}, \forall \mu_1, \mu_2\right\}$. *Then, with this choice of A, the list of vector fields (X, d) satisfies the assumptions of Section* 5.1*, where all of the constants that are admissible (or even m-admissible) in the sense of that section are admissible in the sense of this section. Hence, Theorem* 5.3 *holds for* (X, d).

Proof. We will show that if w_1 and w_2 are words with $\left|\hat{d}(w_1)\right|_1, \left|\hat{d}(w_2)\right|_1 \leq Nd_0$, we have for $\delta \in \mathcal{A}$:

$$\left[\delta^{\hat{d}(w_1)}X_{w_1}, \delta^{\hat{d}(w_2)}X_{w_2}\right] = \sum_{\left|\hat{d}(w_3)\right|_1 \leq Nd_0} c_{w_1,w_2}^{w_3,\delta}\delta^{\hat{d}(w_3)}X_{w_3},$$

with

$$\left\|c_{w_1,w_2}^{w_3,\delta}\right\|_{C^m(\Omega_0)} \lesssim 1.$$

If $\left|\hat{d}(w_1) + \hat{d}(w_2)\right|_1 \leq Nd_0$, this follows easily from the Jacobi identity. We proceed, then, in the case when $\left|\hat{d}(w_1) + \hat{d}(w_2)\right|_1 > Nd_0$. Using that:

$$[X_{w_1}, X_{w_2}] = \sum_{k=1}^{l} c_{w_1,w_2}^{k} X_{w^k}, \tag{5.7}$$

with

$$\left\|c^k_{w_1,w_2}\right\|_{C^m(\Omega_0)} \lesssim 1,$$

and multiplying both sides of (5.7) by:

$$\delta^{\hat{d}(w_1)}\delta^{\hat{d}(w_2)}$$

the result follows easily. ■

Example 5.6. An example to keep in mind where the weakly comparable hypothesis is necessary is given by the following vector fields with formal degrees on $\mathbb{R}^2$:

$$(\partial_x, (1,0,0)), \quad \left(e^{-\frac{1}{x^2}}\partial_y, (0,1,0)\right), \quad (\partial_y, (0,0,1)).$$

If we restrict our attention to the case when $\delta_3 = 0, \delta_1 = \delta_2$ (which is impossible under the weakly comparable hypothesis, without taking $\delta_1 = 0 = \delta_2$) then (without being precise about definitions), we are left with the one-parameter ball of radius δ_1 "generated" by the vector fields:

$$\partial_x, \quad e^{-\frac{1}{x^2}}\partial_y$$

and it is well known that this sort of ball cannot satisfy any sort of doubling condition of the form (5.5).

5.2.2. Multiple lists that span

In this section, we suppose we have ν lists of C^∞ vector fields on $\Omega \subseteq \mathbb{R}^n$ with associated formal degrees:

$$\left(X^\mu_1, d^\mu_1\right), \ldots, \left(X^\mu_{q_\mu}, d^\mu_{q_\mu}\right), d^\mu_j \in (0,\infty), 1 \leq \mu \leq \nu$$

and we assume that for *each* μ, the list

$$X^\mu_1, \ldots, X^\mu_{q_\mu}$$

spans the tangent space at each point in Ω. Our goal is to consider the balls generated by the vector fields:

$$\delta_\mu^{d^\mu_j} X^\mu_j, \quad 1 \leq \mu \leq \nu, \quad 1 \leq j \leq q_\mu,$$

where $\delta = (\delta_1, \ldots, \delta_\nu)$ is small.

Fix $K \Subset \Omega$, a compact subset of Ω, and take $\Omega_0 \Subset \Omega$ such that $K \Subset \Omega_0$. Define:

$$d_0 := \max_{\substack{1\leq\mu\leq\nu \\ 1\leq j\leq q_\mu}} d^\mu_j.$$

We define $\hat{d}_j^\mu \in [0,\infty)^\nu$ for $1 \leq \mu \leq \nu$ and $1 \leq j \leq q_\mu$ to be the vector that is d_j^μ in the μth component and 0 in all the other components. For a list $w = ((w_1,\mu_1),\ldots,(w_r,\mu_r))$ of pairs, where $1 \leq \mu_j \leq \nu$ and $1 \leq w_j \leq q_{\mu_j}$ we define (as in Section 5.2.1):

$$\hat{d}(w) = \sum_{j=1}^{r} \hat{d}_{w_j}^{\mu_j},$$

$$X_w = \mathrm{ad}\left(X_{w_1}^{\mu_1}\right)\mathrm{ad}\left(X_{w_2}^{\mu_2}\right)\cdots\mathrm{ad}\left(X_{w_{r-1}}^{\mu_{r-1}}\right)X_{w_r}^{\mu_r}.$$

Let (X,d) denote the list of vector fields with associated formal degrees given by $\left(X_w, \hat{d}(w)\right)$ where w ranges over all those lists with $\left|\hat{d}(w)\right|_\infty \leq d_0$.

Take $\xi \in (0,1]^\nu$ so small that (X,d) satisfies $\mathcal{C}(x,\xi)$ for every $x \in K$, with Ω_0 taking the place of Ω in the definition of $\mathcal{C}(x,\xi)$.

In this section, we say that C is an admissible constant if C can be chosen to depend only on a fixed upper bound for n, fixed upper and lower bounds for d_j^μ ($1 \leq \mu \leq \nu$, $1 \leq j \leq q_\mu$), a fixed upper bound for ν, a fixed lower bound for:

$$\inf_{\substack{x\in\Omega_0\\ 1\leq\mu\leq\nu}} \left|\det_{n\times n}\left(X_1^\mu(x)|\cdots|X_{q_\mu}^\mu(x)\right)\right|,$$

and fixed upper bounds for a finite number of the norms:

$$\left\|X_j^\mu\right\|_{C^m(\Omega_0)}, \quad 1 \leq \mu \leq \nu, \quad 1 \leq j \leq q_\mu.$$

Theorem 5.7. *Let $\mathcal{A}$ be given by* (5.1)*. Then, with this choice of $\mathcal{A}$, the list of vector fields (X,d) satisfies the assumptions of Section* 5.1*, where all of the constants that are admissible (or even m-admissible) in the sense of that section are admissible in the sense of this section. Hence, Theorem* 5.3 *holds for (X,d).*

Proof. We will show that if w_1 and w_2 are lists with $\left|\hat{d}(w_1)\right|_\infty, \left|\hat{d}(w_2)\right|_\infty \leq d_0$, we have for $\delta \in \mathcal{A}$:

$$\left[\delta^{\hat{d}(w_1)}X_{w_1}, \delta^{\hat{d}(w_2)}X_{w_2}\right] = \sum_{\left|\hat{d}(w_3)\right|_\infty \leq d_0} c_{w_1,w_2}^{w_3,\delta}\delta^{\hat{d}(w_3)}X_{w_3}, \tag{5.8}$$

with

$$\left\|c_{w_1,w_w}^{w_3,\delta}\right\|_{C^m(\Omega_0)} \lesssim 1.$$

If $\left|\hat{d}(w_1)+\hat{d}(w_2)\right|_\infty \leq d_0$, (5.8) follows easily from the Jacobi identity. We proceed, therefore, in the case when $\left|\hat{d}(w_1)+\hat{d}(w_2)\right|_\infty > d_0$. Let us assume that the μth coordinate of $\hat{d}(w_1) + \hat{d}(w_2)$ is greater than d_0. Using that:

$$[X_{w_1}, X_{w_2}] = \sum_{j=1}^{q_\mu} c_{w_1,w_2}^{j} X_j^\mu, \tag{5.9}$$

and multiplying both sides of (5.9) by:

$$\delta^{\hat{d}(w_1)} \delta^{\hat{d}(w_2)}$$

(5.8) follows easily. ∎

Remark 5.8. Theorem 5.7 also follows from the results in Section 4 of [28] (which used the methods of [22]). In fact, the more general results in Section 4 of [28] are clearly a special case Theorem 5.3.

5.2.3. An example where the methods of [22] do not apply

As was already discussed in Section 1.2.1, the methods of [22] fail to prove Theorem 5.3. The main issue is that the error term given by the Campbell-Hausdorff formula cannot be *a priori* controlled using the methods of [22] (see Section 1.2.1). Thus, if one wishes to develop an example where the methods of [22] do not apply, one must use vector fields where the error term is not obviously controllable. As shown in Section 5.2.4 (see also Section 1.2.1), the results of this paper imply that the error term is controllable. The point of this section is to offer an example where the methods of [22] do not prove this fact.

In particular, one needs that the error term of the Campbell-Hausdorff formula not be zero, so the main aspect of the example that follows is that the iterated brackets of the vector fields we present are not eventually zero (this rules out vector fields with polynomial coefficients[24]).

We work in the two-parameter situation, with $\mathcal{A}$ given by (5.1). We consider the list of vector fields on $\mathbb{R}^4$ with formal degrees "generated" by the vector fields

$$(\partial_x + \cos(s)\partial_y, (1,0)), \quad (\partial_s + \cos(x)\partial_t, (0,1)).$$

More specifically, we consider the list of vector fields with formal degrees:

$$(\partial_x + \cos(s)\partial_y, (1,0)), \ (\partial_s + \cos(x)\partial_t, (0,1)) \ (\sin(s)\partial_y - \sin(x)\partial_t, (1,1)),$$
$$(\cos(x)\partial_t, (2,1)), \ (\cos(s)\partial_y, (1,2)), \ (\sin(x)\partial_t, (3,1)), \ (\sin(s)\partial_y, (1,3)).$$

[24] As a consequence, if one is only interested in vector fields with polynomial coefficients, then the methods of [22] (with some adjustments) are sufficient for most purposes.

It is immediate to verify that these vector fields satisfy the assumptions of Theorem 5.3, but (for the reasons mentioned above) the methods of [22] are insufficient to study the balls generated by these vector fields.

5.2.4. Lifting results from the single parameter case and the Campbell-Hausdorff formula

In this section, we discuss a general method whereby one may lift many results from the single parameter setting of [22] to the multi-parameter setting in this paper.

To make this methodology clear, we present a concrete example where it applies. Indeed, this example is interesting in its own right.

We suppose that we are given generating C^∞ vector fields on $\Omega \subseteq \mathbb{R}^n$, with ν parameter formal degrees,

$$(W_1, d_1), \ldots, (W_r, d_r).$$

For a word $w = (w_1, \ldots, w_l)$, $w_j \in \{1, \ldots, r\}$, we define:

$$\hat{d}(w) = \sum_{j=1}^{l} d_{w_j},$$

$$X_w = \mathrm{ad}(X_{w_1}) \cdots \mathrm{ad}\left(X_{w_{l-1}}\right) X_{w_l}.$$

Let $(X, d) = (X_1, d_1), \ldots, (X_q, d_q)$ denote the list of vector fields with formal degrees given by $\left(X_w, \hat{d}(w)\right)$ where $w = (w_1, \ldots, w_l)$ and $l \leq M$ for some fixed large M. Our goal is to show, under the smooth version of the hypotheses of Section 5.1, that the balls

$$B_{\delta^d W}(x)$$

are comparable to the balls

$$B_{(X,d)}(x, \delta).$$

More specificly, fix $x_0 \in \Omega$, and assume (X, d) satisfies $\mathcal{C}(x_0, \xi)$. We assume that we have, for every $\delta \in [0,1)^\nu$ with $\delta \leq \xi$,

$$\left[\delta^{d_i} X_i, \delta^{d_j} X_j\right] = \sum_k c_{i,j}^{k,\delta} \delta^{d_k} X_k,$$

on $B_{(X,d)}(x_0,\delta)$. In what follows, an admissible constant may depend on upper bounds for q and n, lower and upper bounds for the $|\cdot|_1$ norms of the formal degrees, upper bounds for a finite number of the norms $\|X_j\|_{C^m\left(B_{(X,d)}(x_0,\xi)\right)}$ and upper bounds for a finite number of the norms:

$$\sup_{\delta\leq\xi}\sum_{|\alpha|\leq m}\left\|\left(\delta^d X\right)^\alpha c_{i,j}^{k,\delta}\right\|_{C^0\left(B_{(X,d)}(x_0,\delta)\right)},$$

which we assume to be finite–in fact, we only need the above bounds for m which can be chosen to depend only on M and q.

We have,

Theorem 5.9. *There exists an admissible constant $\eta'>0$ such that for every $\delta\leq\xi$, we have:*

$$B_{(X,d)}(x_0,\eta'\delta)\subseteq B_{\delta^d W}(x_0)\subseteq B_{(X,d)}(x_0,\delta).$$

The second containment in Theorem 5.9 is obvious, and so the theorem is really a statement about the first containment. In the single parameter case, Theorem 5.9 was shown in [22]. Specificly, we have:

Theorem 5.10 (Theorem 4 of [22])**.** *In the case $\nu=1$ and when $X_1,\ldots,X_q$ span the tangent space, Theorem 5.9 holds–so long as we allow admissible constants to also depend on a lower bound for:*

$$\left|\det_{n\times n} X(x_0)\right|.$$

Actually, in Theorem 4 of [22], $W_1,\ldots,W_r$ are each given the formal degree 1, but this is not an essential point, and the methods there immediately generalize to give Theorem 5.10. It is worth noting that the proof in [22] uses heavily the Campbell-Hausdorff formula, and therefore use of a lower bound for $|\det_{n\times n} X(x_0)|$ is essential for those methods.

Proof of Theorem 5.9. Apply Theorem 5.3, to obtain Φ_δ, η_1 and ξ_2 as in that theorem. To prove Theorem 5.9, it suffices to construct an admissible constant $\eta'>0$ such that:

$$B_{(X,d)}(x_0,\eta'\delta)\subseteq B_{(\xi_2\delta)^d W}(x_0);$$

rephrasing this, it suffices to show,

$$B_{\left(\delta^d X,\sum d\right)}(x_0,\eta')\subseteq B_{(\xi_2\delta)^d W}(x_0), \tag{5.10}$$

for some admissible $\eta'>0$. Let Y denote the list of vector fields given by the pullback of $\delta^d X$ under the map Φ_δ to $B_{n_0(\delta)}(\eta_1)$, and let W' denote the list of vector fields given by the pullback of $\delta^d W$ under Φ_δ. Pulling back (5.10) via Φ_δ, we see that it suffices to show that,

$$B_{(Y,\sum d)}(0,\eta') \subseteq B_{\xi_2^{\sum d} W'}(0) . \tag{5.11}$$

However, using that the W' generate the Y (since this is just the pullback of the statement that the W generate the X), using that $\left|\det_{n_0(\delta)\times n_0(\delta)} Y(0)\right| \gtrsim 1$ (this follows from (5.3)), and using $\xi_2 \approx 1$, we may apply Theorem 5.10 (in the special case when $\delta \approx 1$) to deduce (5.11), completing the proof. ■

In conclusion, if one can prove a result in the single-parameter setting of [22], one often gets a multi-parameter result "for free," merely by pulling the multi-parameter vector fields back under the scaling map Φ_δ and applying the single-parameter result. In particular, this allows one to use the Campbell-Hausdorff formula to prove results in the multi-parameter setting. This same proof method shows that the error term for the Campbell-Hausdorff formula as discussed in Section 1.2.1 can be controlled in an appropriate sense, even in the multi-parameter setting.

5.3. Control of vector fields

In Section 4.1, we saw that the conditions imposed on the commutators $[X_i, X_j]$ in Section 4 were closely related to three equivalent conditions that were defined in Section 4.1 (see Remark 4.20). The goal in this section is to understand the conditions imposed on the commutators in Section 5.1 in a similar way. To do so, we will lift two of the three equivalent conditions from Section 4.1 into the setting of Section 5.1. These equivalent conditions are interesting in their own right, and will play a role in future work.

We take all the same notation as in Section 5.1, and define (m-)admissible constants in the same way. Let X_{q+1} be a C^1 vector field on Ω, with an associated formal degree $0 \neq d_{q+1} \in [0,\infty)^\nu$. We will introduce conditions on (X_{q+1}, d_{q+1}) which will imply (informally) that one does not "get anything new" if (X_{q+1}, d_{q+1}) is added to the list (X, d). Let $\left(\widehat{X}, \hat{d}\right)$ denote the list of vector fields with formal degrees $(X_1, d_1), \ldots, (X_{q+1}, d_{q+1})$. For an integer $m \geq 1$, we define two conditions (all parameters below are considered to be elements of $(0,\infty)$):

1. $\mathcal{P}_1^m(\kappa_1, \tau_1, \sigma_1, \sigma_1^m)$:
   - $\forall \delta \in \mathcal{A}, x \in K$,
   $$\left| \det_{n_0(x,\delta)\times n_0(x,\delta)} (\delta X)(x) \right|_\infty \geq \kappa_1 \left| \det_{n_0(x,\delta)\times n_0(x,\delta)} \left(\delta \widehat{X}\right)(x) \right|_\infty .$$
   - $\forall x \in K, \left|\det_{j\times j} \widehat{X}(x)\right| = 0$, $n_0(x,\delta) < j \leq n$.

- $\forall \delta \in \mathcal{A}, x \in K, \exists c_{i,q+1}^{j,x,\delta} \in C^0\left(B_{(X,d)}\left(x, \tau_1 \delta\right)\right)$ such that

$$\left[\delta^{d_i} X_i, \delta^{d_{q+1}} X_{q+1}\right] = \sum_{j=1}^{q+1} c_{i,q+1}^{j,x,\delta} \delta^{d_j} X_j, \quad \text{on } B_{(X,d)}\left(x, \tau_1 \delta\right),$$

with

$$\sum_{|\alpha| \leq m-1} \left\| (\delta X)^\alpha c_{i,q+1}^{j,x,\delta} \right\|_{C^0\left(B_{(X,d)}(x,\tau_1\delta)\right)} \leq \sigma_1^m, \quad \left\| c_{i,q+1}^{j,x,\delta} \right\|_{C^0\left(B_{(X,d)}(x,\tau_1\delta)\right)} \leq \sigma_1.$$

2. $\mathcal{P}_3^m\left(\tau_3, \sigma_3, \sigma_3^m\right)$: For every $x \in K, \delta \in \mathcal{A}$, there exist

$$c_j^{x,\delta} \in C^0\left(B_{(X,d)}\left(x_0, \tau_3 \delta\right)\right)$$

such that:

- $\delta^{d_{q+1}} X_{q+1} = \sum_{j=1}^q c_j^{x,\delta} \delta^{d_j} X_j$, on $B_{(X,d)}\left(x, \tau_3 \delta\right)$.
- $\sum_{|\alpha| \leq m} \left\| (\delta X)^\alpha c_j^{x,\delta} \right\|_{C^0\left(B_{(X,d)}(x,\tau_3\delta)\right)} \leq \sigma_3^m$.
- $\sum_{|\alpha| \leq 1} \left\| (\delta X)^\alpha c_j^{x,\delta} \right\|_{C^0\left(B_{(X,d)}(x,\tau_3\delta)\right)} \leq \sigma_3$.

Theorem 5.11. *$\mathcal{P}_1^m \Leftrightarrow \mathcal{P}_3^m$ in the following sense:*

1. *$\mathcal{P}_1^m\left(\kappa_1, \tau_1, \sigma_1, \sigma_1^m\right) \Rightarrow$ there exist admissible constants $\tau_3 = \tau_3(\kappa_1, \tau_1, \sigma_1)$, $\sigma_3 = \sigma_3\left(\kappa_1, \sigma_1\right)$, and an m-admissible constant $\sigma_3^m = \sigma_3^m\left(\kappa_1, \sigma_1^m\right)$ such that $\mathcal{P}_3^m\left(\tau_3, \sigma_3, \sigma_3^m\right)$.*
2. *$\mathcal{P}_3^m(\tau_3, \sigma_3, \sigma_3^m) \Rightarrow$ there exist admissible constants $\kappa_1 = \kappa_1(\sigma_3)$ and $\sigma_1 = \sigma_1(\sigma_3)$ and an m-admissible constant $\sigma_1^m = \sigma_1^m(\sigma_3^m)$, such that $\mathcal{P}_1^m(\kappa_1, \tau_3, \sigma_1, \sigma_1^m)$.*

Furthermore, if $0 < d_{q+1}^\vee$ is a fixed lower bound for $\left|d_{q+1}\right|_1$, then under the condition $\mathcal{P}_3^m\left(\tau_3, \sigma_3, \sigma_3^m\right)$, we have that there exists an admissible constant $\tau' = \tau'\left(d_{q+1}^\vee, \tau_3, \sigma_3\right)$ such that:

$$B_{(X,d)}\left(x, \tau' \delta\right) \subseteq B_{\left(\widehat{X}, \widehat{d}\right)}\left(x, \tau' \delta\right) \subseteq B_{(X,d)}\left(x, \tau_3 \delta\right)$$

for every $x \in K, \delta \in \mathcal{A}$. Finally, if $\eta' \leq \eta_1$ is small enough so that $\Phi_{x,\delta}\left(B_{n_0(x,\delta)}\left(\eta'\right)\right) \subseteq B_{(X,d)}\left(x_0, \tau_3 \delta\right)$ and we define $Y_{q+1}^{x,\delta}$ to be the pullback of $\delta^{d_{q+1}} X_{q+1}$ under $\Phi_{x,\delta}$ to $B_{n_0(x,\delta)}\left(\eta'\right)$, then,

$$\left\| Y_{q+1}^{x,\delta} \right\|_{C^m\left(B_{n_0(x,\delta)}(\eta')\right)} \leq \sigma_4^m$$

where $\sigma_4^m = \sigma_4^m\left(\sigma_3^m\right)$ is an m-admissible constant.

Proof. Merely apply Theorem 4.17 and Propositions 4.18 and 4.19 for each $x \in K, \delta \in A$, to the list of vector fields $(\delta X, \sum d)$, taking $x_0 = x$ and $J_0 = J(x, \delta)$. ∎

Remark 5.12. Our assumption on the commutator $[X_i, X_j]$ in Section 5.1 was essentially that $([X_i, X_j], d_i + d_j)$ satisfied condition $\mathcal{P}_3^m$ for appropriate m.

Definition 5.13. We say a vector field with a formal degree (X_{q+1}, d_{q+1}) is m-controlled by the list of vector fields (X, d) provided either of the two equivalent conditions $\mathcal{P}_1^m$ or $\mathcal{P}_3^m$ holds. We say (X_{q+1}, d_{q+1}) is ∞-controlled by (X, d) if $\mathcal{P}_1^m(\kappa_1, \tau_1, \sigma_1, \sigma_1^m)$ holds for every m, with κ_1, τ_1, and σ_1 independent of m (equivalently if $\mathcal{P}_3^m(\tau_3, \sigma_3, \sigma_3^m)$ holds with τ_3 and σ_3 independent of m).

5.3.1. Examples of control

For this section, we take all the same notation as in Section 5.1, and assume that $\mathcal{A}$ is given by (5.1). As was mentioned in Section 5.3 (see Remark 5.12) our main assumption in Theorem 5.3 is essentially that the commutator $([X_i, X_j], d_i + d_j)$ is "controlled" by (X, d) in the sense of Definition 5.13.

In [22], a stronger assumption was used in the single parameter case (see Section 1.2.1). The most obvious multi-parameter analog of this assumption is the following:

$$[X_i, X_j] = \sum_{d_k \leq d_i + d_j} c_{i,j}^k X_k, \tag{5.12}$$

where the inequality is meant coordinatewise, and the $c_{i,j}^k$ are assumed to be sufficiently smooth. It is easy to see that this assumption is a special case of the assumptions in Section 5.1: indeed, one can take

$$c_{i,j}^{k,x,\delta} := \begin{cases} \delta^{d_i + d_j - d_k} c_{i,j}^k & \text{if } d_k \leq d_i + d_j, \\ 0 & \text{otherwise.} \end{cases}$$

One may wonder whether it is possible to get away with such simple assumptions in applications. This seems to not be the case, and to exemplify the possible difficulties, in this section, we give examples where the closely related notion of control takes a more complicated form.

Example 5.14. This example takes place on $\mathbb{R}^2$ with the vector fields:

$$X_1 = \partial_x, \quad X_2 = x\partial_y, \quad X_3 = \partial_y.$$

Create two copies of these vector fields:

$$X_1^j, X_2^j, X_3^j$$

$j = 1, 2$, both copies acting on the *same* space. We take $\nu = 2$ and assign the formal degrees in $[0,\infty)^2$ as follows:

$$\left(X_1^1, (1,0)\right), \left(X_2^1, (1,0)\right), \left(X_3^1, (2,0)\right), \left(X_1^2, (0,1)\right), \left(X_2^2, (0,1)\right), \left(X_3^2, (0,2)\right).$$

It is clear that:

$$\left(\left[X_1^1, X_2^2\right], (1,0) + (0,1)\right) = (\partial_y, (1,1))$$

is ∞-controlled by the other vector fields. However, it is easy to see that it cannot be written as in (5.12). In this case, one could just throw in the vector field $(\partial_y, (1,1))$, and then the list of vector fields would satisfy (5.12), but this is not the case in the Example 5.16, below. Furthermore, this process of adding in vector fields is counter to the way in which we proceed in Section 6.

Example 5.15. Consider the vector fields with single-parameter formal degrees on $\mathbb{R}$ given by:

$$(\partial_x, 2), \left(x^2\partial_x, 1\right), (x\partial_x, 1.5).$$

Denote them by (X_j, d_j), $j = 1, 2, 3$. We restrict our attention to $|x| \leq 1$. It is clear that for every $|\delta| \leq 1$,

$$\left[\delta^{d_i} X_i, \delta^{d_j} X_j\right] = \sum_k c_{i,j}^{k,\delta} \delta^{d_k} X_k \tag{5.13}$$

with $c_{i,j}^{k,\delta} \in C^\infty$ uniformly in δ. We claim that $(x\partial_x, 1.5)$ is ∞-controlled by the other two vector fields. Indeed, fix x_0, δ, with $\delta, |x_0| \leq 1$. By (5.13) it suffices to show that:

$$\left|\left(\delta^2, x_0^2\delta\right)\right|_\infty \geq \left|\left(\delta^2, x_0^2\delta, x_0\delta^{1.5}\right)\right|_\infty. \tag{5.14}$$

Suppose $\delta^{1.5}|x_0| \geq \delta^2$. Then $|x_0| \geq \sqrt{\delta}$. Hence, $\delta^{1.5}|x_0| \leq \delta |x_0^2|$, completing the proof of (5.14).

What this example shows is that we can write

$$\delta^{1.5} x\partial_x = c_1^{x_0,\delta}\delta^2\partial_x + c_2^{x_0,\delta}\delta x^2\partial_x, \text{ on } B_{(X,d)}(x_0, \tau_3\delta),$$

where τ_3 can be chosen independent of x_0, δ. Note that the choice of $c_1^{x_0,\delta}, c_2^{x_0,\delta}$ depends on x_0 and δ in a way which is more complicated than arises from (5.12): it depends on the ratio of $|x_0|$ and $\sqrt{\delta}$.

Example 5.16. Consider the vector fields with formal degrees on $\mathbb{R}$:

$$(\partial_x, (a, 0)), \quad (\partial_x, (0, b))$$

here $a, b > 0$. In this example, we take $A = \{0 \neq \delta, |\delta| < 1\}$. We will show that the above two vector fields ∞-control $(\partial_x, (c, d))$ if and only if the point (c, d) lies on or above the line going through $(a, 0)$ and $(0, b)$. Here, c, d are any two non-negative real numbers, at least one of which is non-zero.

By replacing $\delta = (\delta_1, \delta_2)$ with $\left(\delta_1^{\frac{1}{a}}, \delta_2^{\frac{1}{b}}\right)$, it is easy to see that it suffices to prove the result for $a = 1 = b$. Hence we need to show that $(\partial_x, (c, d))$ is ∞-controlled by the above two vector fields if and only if $c+d \geq 1$. However, it is easy to see that:

$$(\delta_1, \delta_2)^{(c,d)} \leq C \max\{\delta_1, \delta_2\}$$

for all δ sufficiently small, if and only if $c + d \geq 1$. The result follows easily.

6. Unit operators and maximal functions

In this section, we wish to study maximal operators associated to a special case of the multi-parameter balls from Section 5.1.

Suppose we are given ν families of C^1 vector fields on Ω with associated single-parameter formal degrees:

$$(X^\mu, d^\mu) = \left(\left(X_1^\mu, d_1^\mu\right), \dots, \left(X_{q_\mu}^\mu, d_{q_\mu}^\mu\right)\right), \quad d_j^\mu \in (0, \infty), \quad 1 \leq \mu \leq \nu.$$

We may associate to the X^μs and d^μs a family of vector fields with (multi-parameter) formal degrees. Indeed, let (X, d) denote the list of vector fields X_j^μ, $1 \leq \mu \leq \nu$, $1 \leq j \leq q_\mu$, with the degree of X_j^μ the element of $[0, \infty)^\nu$ which is d_j^μ in the μth coordinate and is 0 in all other coordinates. Define $K \Subset \Omega$ and ξ as in Section 5.2.2, in terms of (X, d). We assume that the list of vector fields (X, d) satisfies all of the assumptions of Section 5.1 (with $\mathcal{A}$ given by (5.1)), *without adding any new vector fields to the list* (X, d).[25] We define admissible constants in the same way as they were defined in Section 5.1. We define, for $\delta \in (0, 1]^\nu$, $\delta \leq \xi$, $x \in K$:

$$B_{(X^1,d^1),\dots,(X^\nu,d^\nu)}(x, \delta) := B_{(X,d)}(x, \delta).$$

We present two interesting examples that satisfy the hypotheses of this section:

[25]In particular, this implies that the (one-parameter) list of vector fields (X^{μ_0}, d^{μ_0}) satisfies the assumptions of Section 5.1, for each μ_0. This can be seen by taking $\delta_\mu = 0$ for every $\mu \neq \mu_0$.

Example 6.1. Suppose $X_1, \ldots, X_m$ are C^∞ vector fields satisfy Hörmander's condition. Use these to generate a list of vector fields with formal degrees (X, d) as in [22] (see Section 1.2.1). Let $\nu = 2$ and let (X, d) be the list of vector fields corresponding to $\mu = 1$. Then let $(\partial_1, 1), \ldots, (\partial_n, 1)$ denote the list of vector fields corresponding to $\mu = 2$. I.e., $\mu = 2$ corresponds to the usual Euclidean vector fields. These then satisfy the hypotheses of the section. The main idea in this example is that one may write:

$$[X_i, \partial_j] = \sum_k a_{i,j}^k \partial_k, \tag{6.1}$$

where the $a_{i,j}^k \in C^\infty$. One must be careful with this example. It is tempting to think, given the results in Section 5.2.2, that one could take any two lists satisfying Hörmander's condition, removing the assumption that one of the lists corresponds to the usual Euclidean vector fields. This is not the case, since the procedure in Section 5.2.2 involved adding more vector fields to the list (X, d), namely the commutators involving vector fields from both lists. In this case, though, this procedure is not necessary, due to (6.1).

Example 6.2. Let (X^μ, d^μ) be ν lists of vector fields with formal degrees such that for each fixed μ, (X^μ, d^μ) satisfies the hypotheses of Section 5.1 (with $\nu = 1$ and $\mathcal{A}$ given by (5.1)). Suppose further that for $\mu_1 \neq \mu_2$, $\left[X_i^{\mu_1}, X_j^{\mu_2}\right] = 0$. Then these vector fields satisfy the hypotheses of this section. In particular, when working on a homogeneous group, one could take $\nu = 2$, and let the $\mu = 1$ vector fields correspond to a homogeneous basis of the left invariant vector fields (with degrees corresponding to their homogeneity) and $\mu = 2$ be a similar list but instead with the right invariant vector fields. This was the setup in [28], and is discussed in Section 1.2.4.

To motivate the results in this section, let us consider a classical example. In this case $\Omega = \mathbb{R}^\nu$, $q_\mu = 1$ for all $1 \leq \mu \leq \nu$, and $(X_1^\mu, d_1^\mu) = (\partial_\mu, 1)$. We have the classical "strong" maximal function in $\mathbb{R}^\nu$. This is given by:

$$\mathcal{M}f(x) := \sup_{\substack{\delta = (\delta_1, \ldots, \delta_n) \\ \delta_j > 0}} \frac{1}{\mathrm{Vol}\left(B_{(\partial_1, 1), \ldots, (\partial_\nu, 1)}(x, \delta)\right)} \int_{B_{(\partial_1, 1), \ldots, (\partial_\nu, 1)}(x, \delta)} |f(z)|\, dz. \tag{6.2}$$

Rewriting (6.2) in the notation of Section 4, we have:

$$\mathcal{M}f(x) = \sup_\delta A_{B_{(\partial_1, 1), \ldots, (\partial_\nu, 1)}(\cdot, \delta)} |f|(x).$$

Perhaps the easiest way to deduce L^p boundedness ($1 < p \leq \infty$) for $\mathcal{M}$ is the idea of Jessen, Marcinkiewicz, and Zygmund [11] to bound $\mathcal{M}$ by a product of the one-dimensional maximal functions, whose L^p boundedness is already understood. To do this, one proves the simple inequality, that there exists a $\lambda > 0$ such that for every δ, and every $f \geq 0$, we have:

$$A_{B_{(\partial_1,1),\ldots,(\partial_\nu,1)}(\cdot,\lambda\delta)} f \leq C A_{B_{(\partial_\nu,1)}(\cdot,\delta_\nu)} \cdots A_{B_{(\partial_1,1)}(\cdot,\delta_1)} f$$

And then it follows immediately, that:

$$\mathcal{M} f(x) \leq C \mathcal{M}_\nu \cdots \mathcal{M}_1 f(x), \tag{6.3}$$

where

$$\mathcal{M}_\mu f(x) = \sup_{\delta_\mu > 0} A_{B_{(\partial_\mu,1)}(\cdot,\delta_\mu)} |f|.$$

In this section, we wish to generalize (6.3).

Theorem 6.3. *There exist admissible constants,* $0 < \tau_2 < \tau_1 < 1$, $\sigma > 0$ *such that for all* $|\delta| \leq \sigma$, *we have, for* $f \in C(\Omega)$, $f \geq 0$, $x \in K$*:*

$$\begin{aligned} A_{B_{(X^1,d^1),\ldots,(X^\nu,d^\nu)}(\cdot,\tau_2\delta)} f(x) &\lesssim A_{B_{(X^\nu,d^\nu)}(\cdot,\tau_1\delta)} \cdots A_{B_{(X^1,d^1)}(\cdot,\tau_1\delta)} f(x) \\ &\lesssim A_{B_{(X^1,d^1),\ldots,(X^\nu,d^\nu)}(\cdot,\delta)} f(x). \end{aligned}$$

Proof. Apply Theorem 4.21 with $(Z^\mu, d^\mu) = (\delta_\mu X^\mu, d^\mu)$, $(X, d) = (\delta X, d)$, and $x_0 = x$, where $x \in K$. We obtain admissible constants $\lambda_1, \lambda_2, \lambda_3$ independent of x, δ as in that theorem.

To conclude the proof, merely take $\tau_1 = \frac{\lambda_2}{\lambda_1}$, $\tau_2 = \frac{\lambda_3}{\lambda_1}$ and replace δ with $\lambda_1 \delta$. ■

Corollary 6.4. *There exist admissible constants* $0 < \tau_2 < \tau_1 < 1$, $\sigma > 0$ *such that if we define, for* $x \in K$, $f \in C(\Omega)$,

$$\mathcal{M} f(x) = \sup_{|\delta| \leq \tau_2 \sigma} A_{B_{(X^1,d^1),\ldots,(X^\nu,d^\nu)}(\cdot,\delta)} |f|(x),$$

and for all $x \in \Omega$ *such that* $B_{(X^\mu,d^\mu)}(x, \tau_1\sigma) \subset \Omega$,

$$\mathcal{M}_\mu f(x) = \sup_{0 < \delta_\mu \leq \tau_1 \sigma} A_{B_{(X^\mu,d^\mu)}(\cdot,\delta_\mu)} |f|(x).$$

Then we have:

$$\mathcal{M} f(x) \lesssim \mathcal{M}_\nu \mathcal{M}_{\nu-1} \cdots \mathcal{M}_1 f(x). \tag{6.4}$$

Proof. This follows directly from Theorem 6.3. ■

Corollary 6.5. *Let* $\mathcal{M}$ *be defined as in Corollary* 6.4. *Then, by possibly admissibly shrinking* τ_2, *we have that* $\mathcal{M}$ *extends to a bounded map* $L^p(\Omega) \to L^p(K)$, *for every* $1 < p < \infty$.

Proof. This would follow from (6.4), provided we have that $\mathcal{M}_1, \ldots, \mathcal{M}_\nu$ extend to bounded operators on L^p $(1 < p < \infty)$. Intuitively, this is simple, since the *one*-parameter balls $B_{(X^\mu, d^\mu)}(\cdot, \delta_\mu)$ satisfy the doubling condition (5.5), and we expect to be able to apply the theory of spaces of homogeneous type to conclude the desired L^p boundedness. There is a slight technicality, though, since if (for a fixed μ), the vector fields X^μ do not span the tangent space, then the balls $B_{(X^\mu, d^\mu)}(\cdot, \delta_\mu)$ do not endow Ω with the structure of a space of homogeneous type: rather, they foliate Ω into leaves, each of which is (locally) a space of homogeneous type. The technical details to deal with this difficulty are covered in Section 6.2. ■

Remark 6.6. Notice, in Corollary 6.5, we have left out $p = \infty$. This is because if the vector fields do not span the tangent space, the maximal operators may only be *a priori* defined on $C(\Omega)$.

Remark 6.7. As mentioned in Section 1.2.4, it is likely that $\mathcal{M}$ is bounded on L^p $(1 < p < \infty)$ for a larger class of balls than is discussed in this section. However, Theorem 6.3 is very tied to the assumptions of this section.

Corollary 6.8. *There exists an admissible constant $\sigma_1 > 0$ such that if for $|\delta| \leq \sigma_1$, we let T_δ denote the Schwartz kernel for the operator:*

$$A_{B_{(X^\nu, d^\nu)}(\cdot, \delta_\nu)} \cdots A_{B_{(X^1, d^1)}(\cdot, \delta_1)},$$

then, for $x \in K$, $T_\delta(x, y)$ is supported on those points (x, y) such that:

$$\inf\left\{\tau > 0 : y \in B_{(X^1, d^1), \ldots, (X^\nu, d^\nu)}(x, \tau\delta)\right\} \lesssim 1, \tag{6.5}$$

and moreover,

$$\sup_y T_\delta(x, y) \approx \frac{1}{\mathrm{Vol}\left(B_{(X^1, d^1), \ldots, (X^\nu, d^\nu)}(x, \delta)\right)}. \tag{6.6}$$

Furthermore, there exists an admissible constant $\sigma_2 > 0$ such that:

$$T_\delta(x, y) \approx \frac{1}{\mathrm{Vol}\left(B_{(X^1, d^1), \ldots, (X^\nu, d^\nu)}(x, \delta)\right)}, \quad y \in B_{(X^1, d^1), \ldots, (X^\nu, d^\nu)}(x, \sigma_2\delta). \tag{6.7}$$

These inequalities are understood to be taking place on the leaf generated by δX passing through x.

Proof. (6.5) and the $\lesssim$ part of (6.6) follow by applying Theorem 6.3 and using the inequality (for $f \geq 0$):

$$A_{B_{(X^\nu, d^\nu)}(\cdot, \tau_1\delta)} \cdots A_{B_{(X^1, d^1)}(\cdot, \tau_1\delta)} f(x) \lesssim A_{B_{(X^1, d^1), \ldots, (X^\nu, d^\nu)}(\cdot, \delta)} f(x)$$

and renaming $\tau_1\delta$, δ. With this new δ, the other half of Theorem 6.3 now reads:

$$A_{B_{(X^1,d^1),\ldots,(X^\nu,d^\nu)}\left(\cdot,\frac{\tau_2}{\tau_1}\delta\right)} f(x) \lesssim A_{B_{(X^\nu,d^\nu)}(\cdot,\delta)} \cdots A_{B_{(X^1,d^1)}(\cdot,\delta)} f(x),$$

thereby establishing (6.7) and therefore the $\gtrsim$ part of (6.6). This completes the proof. ■

Corollary 6.8 has an interesting corollary, to which we now turn. For the statement of this corollary, we restrict our attention to the case $\nu = 2$, but otherwise keep the same assumptions and notation as in the rest of the section.

Corollary 6.9. *Suppose the leaf generated by X^1 passing through x_0 is the same as the leaf generated by X^2 passing through x_0 (call this common leaf L). Then, there exists an admissible constant $\sigma_1 > 0$ such that for every $x_0 \in K$ and every $\delta := (\delta_1, \delta_2)$ with $|\delta| \leq \sigma_1$, we have:*

$$\begin{aligned}\mathrm{Vol}&\left(B_{(X^1,d^1)}(x_0,\delta_1) \cap B_{(X^2,d^2)}(x_0,\delta_2)\right) \\ &\approx \frac{\mathrm{Vol}\left(B_{(X^1,d^1)}(x_0,\delta_1)\right)\mathrm{Vol}\left(B_{(X^2,d^2)}(x_0,\delta_2)\right)}{\mathrm{Vol}\left(B_{(X^1,d^1),(X^2,d^2)}(x_0,\delta)\right)}.\end{aligned}$$

Here, Vol$(\cdot)$ *on the left hand side denotes the induced Lebesgue volume on L.*

Proof. In the following, dy will denote the induced Lebesgue measure on L, and for a set A, χ_A will denote the characteristic function of A.

$$\begin{aligned}\mathrm{Vol}&\left(B_{(X^1,d^1)}(x_0,\delta_1) \cap B_{(X^2,d^2)}(x_0,\delta_2)\right) = \\ &= \int \chi_{B_{(X^1,d^1)}(x_0,\delta_1)}(y)\, \chi_{B_{(X^2,d^2)}(y,\delta_2)}(x_0)\, dy \\ &= \mathrm{Vol}\left(B_{(X^1,d^1)}(x_0,\delta_1)\right)\mathrm{Vol}\left(B_{(X^2,d^2)}(x_0,\delta_2)\right) \\ &\quad\times \int \frac{\chi_{B_{(X^1,d^1)}(x_0,\delta_1)}(y)}{\mathrm{Vol}\left(B_{(X^1,d^1)}(x_0,\delta_1)\right)} \frac{\chi_{B_{(X^2,d^2)}(y,\delta_2)}(x_0)}{\mathrm{Vol}\left(B_{(X^2,d^2)}(x_0,\delta_2)\right)}\, dy.\end{aligned} \tag{6.8}$$

In the above, we have used that $y \in B_{(X^2,d^2)}(x_0,\delta_2)$ if and only if $x_0 \in B_{(X^2,d^2)}(y,\delta_2)$. We use the fact that if $y \in B_{(X^2,d^2)}(x_0,\delta_2)$, then

$$\mathrm{Vol}\left(B_{(X^2,d^2)}(x_0,\delta_2)\right) \approx \mathrm{Vol}\left(B_{(X^2,d^2)}(y,\delta_2)\right),$$

which follows from Theorem 5.3, in particular (5.5). We then have that the RHS of (6.8) is:

$$\begin{aligned}&\approx \mathrm{Vol}\left(B_{(X^1,d^1)}(x_0,\delta_1)\right)\mathrm{Vol}\left(B_{(X^2,d^2)}(x_0,\delta_2)\right) \\ &\quad\times \int \frac{\chi_{B_{(X^1,d^1)}(x_0,\delta_1)}(y)}{\mathrm{Vol}\left(B_{(X^1,d^1)}(x_0,\delta_1)\right)} \frac{\chi_{B_{(X^2,d^2)}(y,\delta_2)}(x_0)}{\mathrm{Vol}\left(B_{(X^2,d^2)}(y,\delta_2)\right)}\, dy \\ &= \mathrm{Vol}\left(B_{(X^1,d^1)}(x_0,\delta_1)\right)\mathrm{Vol}\left(B_{(X^2,d^2)}(x_0,\delta_2)\right) T_\delta(x_0,x_0).\end{aligned}$$

Where T_δ is as in Corollary 6.8. Now the result immediately follows from Corollary 6.8. ■

6.1. Some comments on metrics

Corollary 6.8 has a corollary which can be phrased in terms of metrics, and may serve to give the reader some intuition for these results. We devote this section to this corollary, and maintain all the same notation as in Section 6.

Fix $r = (r_1, \ldots, r_\nu) \in (0,1]^\nu$, and assume $r_\mu = 1$ for some μ. Corresponding to each such r, we obtain a *one*-parameter family of balls:

$$B_{(X,d)}(x, \delta r),$$

for $x \in \Omega$. This one-parameter family of balls is associated to the Carnot-Carathéodory metric ρ_r, associated to the vector fields

$$\left\{ \left(r_\mu^{d_j^\mu} X_j^\mu, d_j^\mu \right) : 1 \leq \mu \leq \nu, 1 \leq j \leq q_\mu \right\}.$$

This metric is defined by:

$$\rho_r(x, y) := \inf \left\{ \delta > 0 : y \in B_{(X,d)}(x, \delta r) \right\}.$$

Remark 6.10. Actually, it could be that ρ_r is not a metric, in that if the X_j^μ do not span the tangent space, the distance between two points might be ∞. This will not affect any of the results in this section.

Remark 6.11. We assumed that $\max_\mu r_\mu = 1$ since, if we drop this assumption, we have:

$$\delta \rho_r = \rho_{\delta^{-1} r}$$

and so every choice of r can be reduced to the case when $\max_\mu r_\mu = 1$.

For each μ we also obtain a metric, the Carnot-Carathéodory metric associated to the vector fields

$$\left\{ \left(X_j^\mu, d_j^\mu \right) : 1 \leq j \leq q_\mu \right\}$$

given by

$$\rho_\mu(x, y) := \inf \left\{ \delta > 0 : y \in B_{(X^\mu, d^\mu)}(x, \delta) \right\},$$

where we have the same caveat as in Remark 6.10.

Given two functions $\Delta_1, \Delta_2 : \Omega \times \Omega \to [0, \infty]$ (one should think of Δ_1, Δ_2 as metrics) we obtain a new function:

$$(\Delta_1 \circ \Delta_2)(x, z) := \inf_{y \in \Omega} \Delta_1(x, y) + \Delta_2(y, z).$$

One should think of $\Delta_1 \circ \Delta_2$ as the "distance" between x and z if one is first allowed to travel in the Δ_1 metric and then in the Δ_2 metric. Of course, even if Δ_1 and Δ_2 are metrics, $\Delta_1 \circ \Delta_2$ may not be symmetric, and therefore will not be a metric.

However, in the case of the ρ_μ above, we do end up with a quasi-metric. Indeed, we have:

Corollary 6.12. *There is an admissible constant $\sigma_2 > 0$ such that for every $x \in K$ and $y \in \Omega$ such that $\rho_r(x,y) < \sigma_2$, we have:*

$$\left[\left(r_1^{-1}\rho_1\right) \circ \left(r_2^{-1}\rho_2\right) \circ \cdots \circ \left(r_\nu^{-1}\rho_\nu\right)\right](x,y) \approx \rho_r(x,y).$$

Proof. $\gtrsim$ is obvious. $\lesssim$ follows from Corollary 6.8. ∎

Remark 6.13. When $\left[X_j^{\mu_1}, X_k^{\mu_2}\right] = 0$ for $\mu_1 \neq \mu_2$ (and with a slight modification in the definition of our metrics), one actually obtains equality in Corollary 6.12 (as opposed to $\approx$). This follows from the proof method in Section 4.1 of [28].

6.2. Foliations whose leaves are locally spaces of homogeneous type

Let $(X_1, d_1), \ldots, (X_q, d_q)$ be vector fields on an open set $\Omega \subseteq \mathbb{R}^n$ with single-parameter formal degrees $d_1, \ldots, d_q \in (0, \infty)$. Under the (single-parameter version of) the hypotheses in Section 5.1, the balls $B_{(X,d)}(x,\delta)$ satisfy the doubling property that is crucial to the theory of spaces of homogeneous type:

$$\operatorname{Vol}\left(B_{(X,d)}(x_0, 2\delta)\right) \lesssim \operatorname{Vol}\left(B_{(X,d)}(x_0, \delta)\right), \tag{6.9}$$

see (5.5). This leads one to consider the maximal operator given by,

$$\mathcal{M}f(x) = \sup_{\delta>0} A_{B_{(X,d)}(\cdot,\delta)} |f|(x) = \sup_{\delta>0} \frac{1}{\operatorname{Vol}\left(B_{(X,d)}(x,\delta)\right)} \int_{B_{(X,d)}(x,\delta)} |f(y)|\, dy,$$

where the supremum is only taken over δ sufficiently small. If the vector fields $X_1, \ldots, X_q$ spanned the tangent space at every point of Ω, the balls $B_{(X,d)}(x,\delta)$ would be open sets of positive Lebesgue measure and (6.9) would imply that they do, in fact, endow Ω with the structure of a space of homogeneous type. Classical arguments then show that $\mathcal{M}$ extends to a bounded operator on L^p $(1 < p \leq \infty)$–this is essentially the situation covered in [22].

However, if the vector fields do not span the tangent space at each point, then the balls $B_{(X,d)}(x,\delta)$ do not turn Ω into space of homogeneous type. Indeed, at a point x_0 where $X_1, \ldots, X_q$ do not span the tangent space, the ball $B_{(X,d)}(x_0,\delta)$ does not even have positive n dimensional Lebesgue measure: it lies on the leaf passing through x_0 generated by $X_1, \ldots, X_q$. This does not prevent $\mathcal{M}$ from extending to a bounded operator on L^p $(1 < p < \infty)$ however, as we shall see. Informally, the idea is that $X_1, \ldots, X_q$ foliate Ω into leaves, where each leaf (endowed with the induced Lebesgue measure) is locally a space of homogeneous type, and the standard theory of maximal functions may be applied. It is crucial, here, that we are working locally (i.e., that we are restricting our attention to $\delta > 0$ small). If we had not restricted our attention to local results, the space of leaves might be quite complicated, to the extent that it would be difficult (if not impossible) to lift the L^p boundedness of $\mathcal{M}$ from each leaf to Ω.[26]

Remark 6.14. Near a non-singular point[27] of the involutive distribution spanned by $X_1, \ldots, X_q$, the boundedness of $\mathcal{M}$ follows immediately. Indeed, in this case, the foliation looks locally like a product space. The maximal function just acts on one of the product variables, and in this variable the balls form a space of homogeneous type. Thus, the main point of this section is to demonstrate an easy way to deal with singular points; however, it will not be necessary for us to make any distinction between non-singular and singular points in our argument.

We now turn to a formal statement of our results. We are given a compact set $K \Subset \Omega$, and $\xi > 0$ such that (X, d) satisfies $\mathcal{C}(x, \xi)$ for every $x \in K$. We assume for every $\delta \leq \xi$, $x \in K$, we have:

$$\left[\delta^{d_i} X_i, \delta^{d_j} X_j\right] = \sum_k c_{i,j}^{k,\delta,x} \delta^{d_k} X_k,$$

on $B_{(X,d)}(x, \delta)$. In addition, we assume:

- The X_js are C^2 on $B_{(X,d)}(x, \xi)$, for every $x \in K$, and satisfy

$$\sup_{x \in K} \|X_j\|_{C^2\left(B_{(X,d)}(x,\xi)\right)} < \infty\,.$$

- For all $|\alpha| \leq 2$, $x \in K$, we have $\left(\delta^d X\right)^\alpha c_{i,j}^{k,\delta,x} \in C^0\left(B_{(X,d)}(x, \delta)\right)$, for every i, j, k, and every $\delta \in \mathcal{A}$, and moreover:

$$\sup_{\substack{\delta \in A \\ x \in K}} \sum_{|\alpha| \leq 2} \left\| \left(\delta^d X\right)^\alpha c_{i,j}^{k,\delta,x} \right\|_{C^0\left(B_{(X,d)}(x,\delta)\right)} < \infty.$$

[26]Consider, for instance, the vector field $\partial_x + \theta \partial_y$ on the manifold $M = \mathbb{R}^2/\mathbb{Z}^2$, where $\theta \in \mathbb{R} \backslash \mathbb{Q}$. In this case, if we denote by $\mathcal{L}$ the space of leaves, we have $L^p(\mathcal{L}) = \mathbb{C}$. Locally, however, M with this foliation just looks like a product space, and the corresponding maximal function is just the standard maximal function along one of the variables.

[27]$x_0 \in \Omega$ is said to be a non-singular point if $\dim \operatorname{span}\{X_1(x), \ldots, X_q(x)\}$ is constant in a neighborhood of x_0.

Finally, let

$$n_0(x,\delta) = \dim \operatorname{span}\left\{\delta^{d_1} X_1(x), \ldots, \delta^{d_q} X_q(x)\right\}.$$

We say C is an admissible constant if C can be chosen to depend only on fixed upper and lower bounds $d_{max} < \infty$, $d_{min} > 0$, for $d_1, \ldots, d_q$, a fixed upper bound for n, q and a fixed upper bound for the quantities:

$$\sup_{x \in K} \|X_j\|_{C^2\left(B_{(X,d)}(x,\xi)\right)}, \quad \sup_{\substack{\delta \in A \\ x \in K}} \sum_{|\alpha| \leq 2} \left\| \left(\delta^d X\right)^\alpha c_{i,j}^{k,\delta,x} \right\|_{C^0\left(B_{(X,d)}(x,\delta)\right)}.$$

Theorem 6.15. *There exists an admissible constant $\xi' > 0$, $\xi' \leq \xi$, such that if we define, for $f \in C(\Omega)$, $x \in K$,*

$$\mathcal{M}f(x) = \sup_{0 < \delta \leq \xi'} A_{B_{(X,d)}(\cdot,\delta)} |f|(x), \tag{6.10}$$

then for every $f \in C(\Omega) \cap L^p(\Omega)$,

$$\|\mathcal{M}f\|_{L^p(K)} \leq C_p \|f\|_{L^p(\Omega)},$$

for every $1 < p \leq \infty$. Here, C_p is an admissible constant which may also depend on p.

The main assumptions of this section are equivalent to saying that Theorem 5.3 applies to the vector fields (X,d). Let ξ_1, η_1 be admissible constants as in the conclusions of Theorem 5.3. Define,

$$\Omega' = \bigcup_{x \in K} B_{(X,d)}\left(x, \frac{\xi_1}{2}\right), \qquad \Omega'' = \bigcup_{x \in K} B_{(X,d)}\left(x, \frac{\xi_1}{4}\right).$$

Theorem 6.15 will follow from the following two propositions.

Proposition 6.16. *There exists an admissible constant $\xi_0 > 0$, $\xi_0 < \xi$, such that for every $\xi' \leq \xi_0$ and every $f \in C(\Omega)$ with $f \geq 0$, we have:*

$$\int_K f(x)\, dx \lesssim \int_{\Omega''} A_{B_{(X,d)}(\cdot,\xi')} f(x)\, dx \lesssim \int_{\Omega'} f(x)\, dx,$$

where the implicit constants are admissible but also allowed to depend on a lower bound for ξ'.

Proposition 6.17. $\xi' < \frac{\xi_1}{2}$, *We have the pointwise bound, for $1 < p \leq \infty$, $0 < \xi' \leq \frac{\xi_1}{4}$, and $x \in \Omega''$:*

$$A_{B_{(X,d)}(\cdot,\xi')} |\mathcal{M}f|^p(x) \lesssim A_{B_{(X,d)}(\cdot,2\xi')} |f|^p(x),$$

where the implicit constant is admissible and can also depend on p and a lower bound for ξ', but not on x.

Proof of Theorem 6.15 given Propositions 6.16 and 6.17. Fix $p > 1$. Let ξ_0 be as in Propositions 6.16. Fix $\xi' > 0$ an admissible constant, such that $\xi' < \min\left\{\frac{\xi_0}{2}, \frac{\xi_1}{4}\right\}$. Take $f \in C(\Omega) \cap L^p(\Omega)$, and consider:

$$\begin{aligned}\|\mathcal{M}f\|_{L^p(K)}^p &= \int_K (\mathcal{M}f(x))^p \, dx \lesssim \int_{\Omega''} A_{B_{(X,d)}(\cdot,\xi')} |\mathcal{M}f|^p(x) \, dx \\ &\lesssim \int_{\Omega''} A_{B_{(X,d)}(\cdot,2\xi')} |f|^p(x) \, dx \lesssim \int_{\Omega'} |f(x)|^p \, dx \lesssim \|f\|_{L^p(\Omega)}^p ,\end{aligned}$$

completing the proof. ■

We now prove Proposition 6.16. To do so, we need two lemmas.

Lemma 6.18. *There exists an admissible constant $\eta^0 > 0$ such that for every $0 < \eta' \leq \eta^0$, and every $f \in C(\Omega)$ with $f \geq 0$, we have:*

$$\int_K f(x) \, dx \lesssim \frac{1}{(2\eta')^q} \int_{|t| \leq \eta'} \int_{\Omega''} f\left(e^{t\cdot X}x\right) \, dx \, dt \lesssim \int_{\Omega'} f(x) \, dx.$$

Proof. Note, for $|t| \leq \frac{\xi_1}{4}$, we have

$$\Omega' \supseteq e^{t\cdot X}\Omega'' \supseteq K.$$

To make use of this, we choose $\eta^0 \leq \frac{\xi_1}{4}$. Furthermore, by taking $\eta^0 > 0$ admissible small enough, we have for all $|t| \leq \eta^0$, and all $x \in \Omega''$,

$$\left|\det \frac{\partial}{\partial_x} e^{t\cdot X}x\right| \geq \frac{1}{2}.$$

This follows since the C^2 norm of $e^{t\cdot X}x - x$ is admissibly bounded (see Theorem A.1) and because when $t = 0$, $e^{t\cdot X}x = x$.

Putting these results together, we have from a simple change of variables, for $|t| \leq \eta^0$,

$$\int_K f(x) \, dx \lesssim \int_{\Omega''} f\left(e^{t\cdot X}x\right) \, dx \lesssim \int_{\Omega'} f(x) \, dx.$$

Averaging both sides over $|t| \leq \eta'$ yields the proof. ■

Lemma 6.19. *Let η^0 be as in Lemma 6.18. There exists an admissible constant $\xi_0 > 0$ such that for every $0 < \xi' \leq \xi_0$, there exist admissible constants $0 < \eta'' = \eta''(\xi')$, $0 < \eta'$, $\eta'', \eta' \leq \eta^0$ such that*[28] *for every $f \in C(\Omega)$, $f \geq 0$, we have:*

$$\int_{|t| \leq \eta''} f\left(e^{t\cdot X}x\right) \, dt \lesssim A_{B_{(X,d)}(\cdot,\xi')} f(x) \lesssim \int_{|t| \leq \eta'} f\left(e^{t\cdot X}x\right) \, dt,$$

for every $x \in \Omega'$. Here, the implicit constants are allowed to depend on a lower bound for ξ'.

[28]Here, η'' can be chosen to depend only on a fixed lower bound for ξ'.

Proof. This follows just as in Proposition 4.22. The straightforward modifications are left to the reader. ■

Proof of Proposition 6.16. Take η^0, ξ_0 as in Lemmas 6.18 and 6.19. For $\xi' \leq \xi_0$, let η', η'' be as in the conclusion of Lemma 6.19 (here and in the rest of the proof, all constants are allowed to depend on a lower bound for ξ'–so that, in particular, $\eta'' \gtrsim 1$). We then have, using Lemmas 6.18 and 6.19 freely, for $f \in C(\Omega)$, with $f \geq 0$,

$$\begin{aligned}\int_K f(x)\, dx &\lesssim \int_{|t|\leq\eta''}\int_{\Omega''} f\left(e^{t\cdot X}x\right)\, dx\, dt \lesssim \int_{\Omega''} A_{B_{(X,d)}(\cdot,\xi')} f(x)\, dx\\ &\lesssim \int_{|t|\leq\eta'}\int_{\Omega''} f\left(e^{t\cdot X}x\right)\, dx\, dt \lesssim \int_{\Omega'} f(x)\, dx,\end{aligned}$$

which completes the proof. ■

Proof of Proposition 6.17. Fix $x \in \Omega''$ and ξ' as in the statement of the proposition. In what follows all implicit admissible constants are also allowed to depend on a lower bound for ξ'. We define the maximal function $\mathcal{M}$ in terms of this fixed ξ', as in (6.10). By definition of Ω'', there exists $x_0 \in K$ such that $B_{(X,d)}(x,\xi') \subseteq B_{(X,d)}(x_0,\xi_1)$. Let $n_0 = \dim \operatorname{span}\{X_1(x_0), \ldots, X_q(x_0)\}$, and let $\Phi : B_{n_0}(\eta_1) \to B_{(X,d)}(x_0,\xi)$ be the map guaranteed by Theorem 5.3 where we take $\delta = \xi$ and $x = x_0$. Note that,

$$B_{(X,d)}(x,\xi') \subseteq B_{(X,d)}(x_0,\xi_1) \subseteq \Phi\left(B_{n_0}(\eta_1)\right).$$

Let $Y_1, \ldots, Y_q$ be the pullbacks of $X_1, \ldots, X_q$ via the map Φ, to $B_{n_0}(\eta)$. Note, for $u \in B_{n_0}(\eta)$ and $\delta > 0$ small enough that $B_{(Y,d)}(u,\delta) \Subset B_{n_0}(\eta)$, we have

$$\Phi\left(B_{(Y,d)}(u,\delta)\right) = B_{(X,d)}(\Phi(u),\delta).$$

Using that $|\det_{n_0\times n_0} d\Phi(u)| \approx |\det_{n_0\times n_0} X(x_0)|$, and applying a change of variables as in (B.2), we see that

$$\operatorname{Vol}\left(B_{(X,d)}(\Phi(u),\delta)\right) \approx \left|\det_{n_0\times n_0} X(x_0)\right| \operatorname{Vol}\left(B_{(Y,d)}(u,\delta)\right).$$

It follows, for $f \in C(\Omega)$, with $f \geq 0$,

$$\begin{aligned}A_{B_{(X,d)}(\cdot,\delta)}(f\circ\Phi^{-1})(\Phi(u)) &= \frac{1}{\operatorname{Vol}(B_{(X,d)}(\Phi(u),\delta))}\int_{B_{(X,d)}(\Phi(u),\delta)} f\left(\Phi^{-1}(y)\right) dy\\ &\approx \frac{1}{\operatorname{Vol}\left(B_{(Y,d)}(u,\delta)\right)}\int_{B_{(Y,d)}(u,\delta)} f(v)\, dv\\ &= A_{B_{(Y,d)}(\cdot,\delta)} f(u),\end{aligned}$$

where, again, we have used (B.2) and dv denotes Lebesgue measure on $B_{n_0}(\eta)$ and dy denotes the induced Lebesgue measure on the leaf in which $B_{(X,d)}(\Phi(u),\delta)$ lies. Consider, for $y\in B_{(X,d)}\left(x_0,\frac{\xi_1}{2}\right)$,

$$\begin{aligned}\mathcal{M}\left(f\circ\Phi^{-1}\right)(y) &= \sup_{\xi'\geq\delta>0} A_{B_{(X,d)}(\cdot,\delta)}\left|f\circ\Phi^{-1}\right|(y)\\ &\approx \sup_{\xi'\geq\delta>0} A_{B_{(Y,d)}(\cdot,\delta)}|f|\left(\Phi^{-1}(y)\right) =: \widetilde{\mathcal{M}}f\left(\Phi^{-1}(y)\right),\end{aligned}$$

where $\widetilde{\mathcal{M}}$ denotes the maximal function defined in terms of the Carnot-Carathéodory balls defined by the vector fields (Y,d).

Hence, we have,

$$\begin{aligned}A_{B_{(X,d)}(\cdot,\xi')}\left|\mathcal{M}\left(f\circ\Phi^{-1}\right)\right|^p(x) &\approx A_{B_{(X,d)}(\cdot,\xi')}\left[\left|\widetilde{\mathcal{M}}f\right|^p\circ\Phi^{-1}\right](x)\\ &\approx A_{B_{(Y,d)}(\cdot,\xi)}\left|\widetilde{\mathcal{M}}f\right|^p\left(\Phi^{-1}(x)\right).\end{aligned}$$

Similarly, we have

$$A_{B_{(X,d)}(\cdot,2\xi')}\left|f\circ\Phi^{-1}\right|^p(x)\approx A_{B_{(Y,d)}(\cdot,2\xi')}|f|^p\left(\Phi^{-1}(x)\right).$$

Thus, to complete the proof, it suffices to show the bound

$$A_{B_{(Y,d)}(\cdot,\xi')}\left|\widetilde{\mathcal{M}}f\right|^p\left(\Phi^{-1}(x)\right)\lesssim A_{B_{(Y,d)}(\cdot,2\xi')}|f|^p\left(\Phi^{-1}(x)\right).$$

Moreover, since $\mathrm{Vol}\left(B_{(Y,d)}\left(\Phi^{-1}(x),\xi'\right)\right)\approx\mathrm{Vol}\left(B_{(Y,d)}\left(\Phi^{-1}(x),2\xi'\right)\right)$ (in fact both are ≈ 1, but we will not need this), it suffices to show

$$\left\|\widetilde{\mathcal{M}}f\right\|^p_{L^p\left(B_{(Y,d)}(\Phi^{-1}(x),\xi')\right)}\lesssim\|f\|^p_{L^p\left(B_{(Y,d)}(\Phi^{-1}(x),2\xi')\right)}.$$

This is immediate from the classical theory of spaces of homogeneous type, since the balls $B_{(Y,d)}(\cdot,\delta)$ satisfy all the axioms of a space of homogeneous type, uniformly in the relevant parameters. ■

A. Two results from calculus

In this appendix, we discuss two theorems from calculus that we will use throughout the paper: a uniform version of the inverse function theorem, and how the smoothness of $e^{t_1X_1+t_2X_2+\cdots+t_\nu X_\nu}x_0$, as a function of $t_1,\dots,t_\nu$, depends on the smoothness of $X_1,\dots,X_\nu$. These results are surely familiar, in some form or another, to the reader. However, they play such a fundamental role in our analysis, that we feel it is prudent to state them in the precise form we shall use them.

For a C^1 vector field Y, one defines $E(t) = e^{tY}x_0$ to be the unique solution to the ODE $\frac{d}{dt}E(t) = Y(E(t))$ satisfying $E(0) = x_0$. This unique solution always exists for $|t|$ sufficiently small (depending on the C^1 norm of Y). This allows us to define:

$$e^{t_1 X_1 + \cdots + t_\nu X_\nu} x_0$$

for $|t|$ sufficiently small, where $t = (t_1, \ldots, t_\nu)$. We have:

Theorem A.1. *Suppose $X_1, \ldots, X_\nu$ are C^m vector fields ($m \geq 1$), defined on an open set $\Omega \subseteq \mathbb{R}^n$. Then, for x_0 fixed, the function:*

$$u(t) = e^{t_1 X_1 + \cdots + t_\nu X_\nu} x_0 - x_0$$

is C^m. Moreover, the C^m norm of this function can be bounded in terms of n, ν and the C^m norms of $X_1, \ldots, X_\nu$.

Proof. It is perhaps easiest to consider the function:

$$v(\epsilon, t) = e^{\epsilon(t_1 X_1 + \cdots + t_\nu X_\nu)} x_0.$$

Then, $u(t) = v(1,t) - x_0$, and v is defined by an ODE in the ϵ variable. That $v - x_0$ is C^m (in both variables) is classical. See [5, Chapter X]. Alternatively, one can modify the proof method in [10] to this situation for a more elegant proof. ■

Remark A.2. One could write t in polar coordinates r, ω, and consider the function, $f(r,\omega) = u(r\omega)$. Then, one has, $f \in C^m(r,\omega)$. Moreover, one has, for $a + |b| = m$ ($a \in \mathbb{N}$, b a multi-index),

$$\partial_r \partial_r^a \partial_\omega^b f(r,\omega)$$

exists and is continuous.

We now turn to the inverse function theorem:

Theorem A.3. *Fix an open set $U \subseteq R^n$, and fix $x_0 \in U$. Suppose $K \subset C^1(U; R^n)$ is a compact set such that for all $f \in K$, $\det df(x_0) \neq 0$, and hence, $|\det df(x_0)|$ is bounded away from 0 uniformly for $f \in K$. Then, there exist constants $\delta_1, \delta_2 > 0$, such that for all $f \in K$,*

- *$f|_{B(x_0,\delta_1)}$ is a C^1 diffeomorphism onto its image.*
- *$B(f(x_0), \delta_2) \subseteq f(B(x_0, \delta_1))$.*

here, $B(x_0, \delta)$ denotes the usual Euclidean ball centered at x_0 of radius δ.

Proof. This follows from a straight-forward modification of the proof in [26], by using the Arzelà-Ascoli theorem. Alternatively, since in our proofs, we will always show that the relevant set is a pre-compact subset of C^1, by showing it is a bounded subset of C^2, the derivatives of the functions in our set will actually be uniformly Lipschitz, and in this case one may use the theorem in [9]. ■

B. The Cauchy-Binet formula

In this appendix, we review the Cauchy-Binet formula and an associated change of variables formula, that is essential to the work in this paper. We first recall some notation from Section 1.1: given two integers $1 \leq m \leq n$, define $\mathcal{I}(m,n)$ to be the set of all lists of integers $(i_1, \ldots, i_m)$, such that:

$$1 \leq i_1 < i_2 < \cdots < i_m \leq n.$$

Given an $n \times q$ matrix A, and $n_0 \leq n \wedge q$, for $I \in \mathcal{I}(n_0, n)$, $J \in \mathcal{I}(n_0, q)$ define the $n_0 \times n_0$ matrix $A_{I,J}$ by using the rows from A which are listed in I and the columns from A listed in J. We define:

$$\det_{n_0 \times n_0} A = \left(\det A_{I,J}\right)_{\substack{I \in \mathcal{I}(n_0,n) \\ J \in \mathcal{I}(n_0,q)}}.$$

In particular, $\det_{n_0 \times n_0} A$ is a vector (it will not be important to us in which order the coordinates are arranged). A special case arises when $q = n_0$. Indeed, in this case, we have ([29, p.127]):

$$\left|\det_{n_0 \times n_0} A\right| = \sqrt{\det A^t A} \tag{B.1}$$

and both of these quantities are equal to the volume of the n_0 dimensional parallelepiped with edges given by the columns of A. This is a special case of the Cauchy-Binet formula. Because of this, we obtain a change of variables formula which will be of use to us. Suppose Φ is a C^1 diffeomorphism from an open subset U in $\mathbb{R}^{n_0}$ mapping to an n_0 dimensional submanifold of $\mathbb{R}^n$, where this submanifold is given the Lebesgue measure, dx. Then, we have:

$$\int_{\Phi(U)} f(x)\, dx = \int_U f(\Phi(t)) \left|\det_{n_0 \times n_0} (d\Phi(t))\right| dt. \tag{B.2}$$

References

[1] Bramanti, M., Brandolini, L. and Pedroni, M.: Basic properties of nonsmooth Hörmander's vector fields and Poincaré's inequality. Preprint available at `arXiv:0809.2872`, 2008.

[2] Chevalley, C.: *Theory of Lie Groups. I.* Princeton Mathematical Series **8**. Princeton University Press, Princeton, NJ, 1946.

[3] Christ, M.: Regularity properties of the $\overline{\partial}_b$ equation on weakly pseudoconvex CR manifolds of dimension 3. *J. Amer. Math. Soc.* **1** (1988), no. 3, 587–646.

[4] Chang, D.-C., Nagel, A. and Stein, E. M.: Estimates for the $\overline{\partial}$-Neumann problem in pseudoconvex domains of finite type in $\mathbb{C}^2$. *Acta Math.* **169** (1992), no. 3-4, 153–228.

[5] Dieudonné, J.: *Foundations of modern analysis*. Pure and Applied Mathematics **10**. Academic Press, New York, 1960.

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

[7] Fefferman, C. L. and Sánchez-Calle, A.: Fundamental solutions for second order subelliptic operators. *Ann. of Math. (2)* **124** (1986), no. 2, 247–272.

[8] Hermann, R.: The differential geometry of foliations. II. *J. Math. Mech.* **11** (1962), 303–315.

[9] Hubbard, J. H. and Hubbard, B. B.: *Vector calculus, linear algebra, and differential forms. A unified approach.* Prentice Hall, Upper Saddle River, NJ, 1999.

[10] Izzo, A. J.: C^r convergence of Picard's successive approximations. *Proc. Amer. Math. Soc.* **127** (1999), no. 7, 2059–2063.

[11] Jessen, B., Marcinkiewicz, J. and Zygmund, A.: Note on the differentiability of multiple integrals. *Funda. Math.* **25** (1935), 217–234.

[12] Jerison, D. and Sánchez-Calle, A.: Subelliptic, second order differential operators. In *Complex analysis, III (College Park, Md., 1985–86)*, 46–77. Lecture Notes in Math. **1277**. Springer, Berlin, 1987.

[13] Koenig, K. D.: On maximal Sobolev and Hölder estimates for the tangential Cauchy-Riemann operator and boundary Laplacian. *Amer. J. Math.* **124** (2002), no. 1, 129–197.

[14] Lundell, A. T.: A short proof of the Frobenius theorem. *Proc. Amer. Math. Soc.* **116** (1992), no. 4, 1131–1133.

[15] Montanari, A. and Morbidelli, D.: Nonsmooth Hörmander's vector fields and their control balls. To appear in *Trans. Amer. Math. Soc.*

[16] Müller, D., Ricci, F. and Stein, E. M.: Marcinkiewicz multipliers and multi-parameter structure on Heisenberg (-type) groups. I. *Invent. Math.* **119** (1995), no. 2, 199–233.

[17] Nagel, A., Ricci, F. and Stein, E. M.: Singular integrals with flag kernels and analysis on quadratic CR manifolds. *J. Funct. Anal.* **181** (2001), no. 1, 29–118.

[18] Nagel, A., Rosay, J.-P., Stein, E. M. and Wainger, S.: Estimates for the Bergman and Szegő kernels in $\mathbf{C}^2$. *Ann. of Math. (2)* **129** (1989), no. 1, 113–149.

[19] Nagel, A. and Stein, E. M.: Differentiable control metrics and scaled bump functions. *J. Differential Geom.* **57** (2001), no. 3, 465–492.

[20] Nagel, A. and Stein, E. M.: On the product theory of singular integrals. *Rev. Mat. Iberoamericana* **20** (2004), no. 2, 531–561.

[21] Nagel, A. and Stein, E. M.: The $\overline{\partial}_b$-complex on decoupled boundaries in $\mathbb{C}^n$. *Ann. of Math. (2)* **164** (2006), no. 2, 649–713.

[22] Nagel, A., Stein, E. M. and Wainger, S.: Balls and metrics defined by vector fields. I. Basic properties. *Acta Math.* **155** (1985), no. 1-2, 103–147.

[23] Rampazzo, F.: Frobenius-type theorems for Lipschitz distributions. *J. Differential Equations* **243** (2007), no. 2, 270–300.

[24] Rothschild, L. P. and Stein, E. M.: Hypoelliptic differential operators and nilpotent groups. *Acta Math.* **137** (1976), no. 3-4, 247–320.

[25] Sánchez-Calle, A.: Fundamental solutions and geometry of the sum of squares of vector fields. *Invent. Math.* **78** (1984), no. 1, 143–160.

[26] Spivak, M.: *Calculus on manifolds. A modern approach to classical theorems of advanced calculus*. W. A. Benjamin, New York-Amsterdam, 1965.

[27] Stein, E. M.: *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*. Princeton Mathematical Series **43**. Monographs in Harmonic Analysis, III. Princeton University Press, Princeton, NJ, 1993.

[28] Street, B.: An algebra containing the two-sided convolution operators. *Adv. Math.* **219** (2008), no. 1, 251–315.

[29] Thrall, R. M. and Tornheim, L.: *Vector spaces and matrices*. John Wiley & Sons, New York, 1957.

[30] Tao, T. and Wright, J.: L^p improving bounds for averages along curves. *J. Amer. Math. Soc.* **16** (2003), no. 3, 605–638 (electronic).

Recibido: 11 de mayo de 2009
Revisado: 10 de noviembre de 2009

Brian Street
University of Wisconsin-Madison
Department of Mathematics
480 Lincoln Dr.
Madison, WI, 53706, USA
street@math.wisc.edu

The author was partially supported by NSF DMS-0802587.