

COORDINATES ADAPTED TO VECTOR FIELDS: CANONICAL COORDINATES

BETSY STOVALL AND BRIAN STREET



Abstract. Given a finite collection of C^1 vector fields on a C^2 manifold which span the tangent space at every point, we consider the question of when there is locally a coordinate system in which these vector fields have a higher level of smoothness. For example, when is there a coordinate system in which the vector fields are smooth, or real analytic, or have Zygmund regularity of some finite order? We address this question in a quantitative way, which strengthens and generalizes previous works on the quantitative theory of sub-Riemannian (aka Carnot–Carathéodory) geometry due to Nagel, Stein, and Wainger, Tao and Wright, the second author, and others. Furthermore, we provide a diffeomorphism invariant version of these theories. This is the first part in a three part series of papers. In this paper, we study a particular coordinate system adapted to a collection of vector fields (sometimes called canonical coordinates) and present results related to the above questions which are not quite sharp; these results form the backbone of the series. The methods of this paper are based on techniques from ODEs. In the second paper, we use additional methods from PDEs to obtain the sharp results. In the third paper, we prove results concerning real analyticity and use methods from ODEs.

Contents

1	Introduction	1781
2	Function Spaces	1784
2.1	Function spaces on Euclidean space.	1784
2.2	Function spaces on manifolds.	1785
2.2.1	Beyond manifolds.	1788
3	Overview of the Series	1788
3.1	Qualitative results.	1788
3.2	Quantitative results.	1790
3.2.1	Diffeomorphism invariance.	1792
4	Main Results of this Paper	1793
4.1	More on the assumptions.	1796
5	Wedge Products	1797
6	Densities	1798
7	Scaling and Other Consequences	1800

7.1	Classical sub-Riemmanian geometries and the work of Nagel, Stein, and Wainger.	1801
7.1.1	Hörmander's condition.	1802
7.2	Multi-parameter balls.	1804
7.3	Generalized sub-Riemannian geometries.	1805
7.4	Diffeomorphism invariance and nonsmooth vector fields.	1809
7.5	Several complex variables.	1810
8	Function Spaces, Revisited	1811
8.1	Comparison with Euclidean function spaces.	1815
9	Proofs	1818
9.1	An ODE.	1818
9.1.1	Derivation of the ODE.	1818
9.1.2	Regularity properties.	1821
9.2	An inverse function theorem.	1829
9.3	Proof of the main result.	1832
9.3.1	Linearly independent.	1832
9.3.2	Linearly dependent.	1838
9.4	Densities.	1846
9.5	More on the assumptions.	1852
A	Proof of Proposition 3.1	1852
	References	1860

1 Introduction

Let $X_1, \dots, X_q$ be C^1 vector fields on a C^2 manifold M , which span the tangent space at every point of M . For $s > 0$ let $\mathcal{C}^s$ denote the Zygmund space of order s (see Section 2.1), let $\mathcal{C}^\infty$ denote the space of smooth functions, and let $\mathcal{C}^\omega$ denote the space of real analytic functions. In this three part series of papers, we investigate the following closely related questions for $s \in (1, \infty] \cup \{\omega\}$:¹

- (i) When is there a coordinate system near a fixed point $x_0 \in M$ such that the vector fields $X_1, \dots, X_q$ are $\mathcal{C}^{s+1}$ in this coordinate system?
- (ii) When is there a $\mathcal{C}^{s+2}$ manifold structure on M , compatible with its C^2 structure, such that $X_1, \dots, X_q$ are $\mathcal{C}^{s+1}$ with respect to this structure? When such a structure exists, we will see it is unique.
- (iii) When there is a coordinate system as in (i), how can we pick it so that $X_1, \dots, X_q$ are “normalized” in this coordinate system in a quantitative way which is useful for applying techniques from analysis?

We present necessary and sufficient, coordinate free, conditions for (i) and (ii) and, under these conditions, give a quantitative answer to (iii). See Section 3 for an overview of the results of this series. The outline of this series is as follows:

- (I) In this paper, we study a particular explicit coordinate system adapted to a collection of vector fields. This coordinate system is sometimes known (at least in the setting of Lie groups) as canonical coordinates of the first kind.

¹ We define $\infty + 1 = \infty + 2 = \infty$ and $\omega + 1 = \omega + 2 = \omega$.

This builds on previous work of Nagel, Stein, and Wainger [NSW85], Tao and Wright [TW03, Section 4], and the second author [Str11]. To study these canonical coordinates, we use methods from ODEs. Unfortunately, the results given by these methods are one derivative short of being optimal (see Remark 4.8).

- (II) In the second paper [Str18a], we obtain the optimal results (in terms of Zygmund spaces) by introducing a new (implicitly defined) coordinate system. The second paper takes as a starting point the main result of this paper, and then uses methods from PDEs to obtain the sharp results. These PDE methods were inspired by, and are closely related to, Malgrange’s celebrated proof of the Newlander–Nirenberg theorem [Mal69].
- (III) While the second paper obtains optimal regularity in terms of Zygmund spaces, the methods there are not applicable to the real analytic setting. In the third paper [Str18b], we return to canonical coordinates and methods from ODEs to obtain results regarding real analyticity. The third paper takes the main results of this paper as its starting point.

To help explain the sorts of questions we investigate, we consider a trivial example.

EXAMPLE 1.1. Let X be a C^1 vector field on a C^2 manifold $\mathfrak{M}$ with $X(x_0) \neq 0$ for some $x_0 \in \mathfrak{M}$. Let M be the integral curve of X passing through x_0 . It is well known that there is a unique C^2 manifold structure on M which sees $M \hookrightarrow \mathfrak{M}$ as a C^2 injective immersion (see Proposition 3.1); X spans the tangent space to M at every point. Set $\Phi(t) := e^{tX}x_0$ and let $I \subseteq \mathbb{R}$ be a maximal open interval containing 0 such that Φ is defined on I and $\Phi : I \rightarrow M$ is injective. It is easy to see that $\Phi|_I$ is a C^2 diffeomorphism onto its image, and therefore Φ defines a coordinate chart on M near x_0 . In this coordinate system X equals $\frac{\partial}{\partial t}$; more precisely, $\Phi^*X = \frac{\partial}{\partial t}$. Thus, we have not only picked a coordinate chart on M in which X is smooth, but we have also chosen it so that X is “normalized” to be $\frac{\partial}{\partial t}$.

It is straightforward to generalize Example 1.1 to a finite collection of vector fields, so long as the vector fields are assumed to commute. The purpose of this series of papers is to consider similar results when the vector fields are not assumed to commute; in which case it is not always possible to pick a coordinate system in which the vector fields are smooth. Indeed, we present necessary and sufficient conditions for when one can pick a coordinate system giving the vector fields a desired level of regularity.

The coordinate charts developed in this series can be viewed as scaling maps in a wide variety of problems; this is described in more detail in Section 7. Seen in this perspective, these results are the latest, most general, and sharpest in a series of papers on the quantitative theory of sub-Riemannian (or Carnot–Carathéodory) geometry. This started with the foundational work of Nagel, Stein, and Wainger [NSW85] and the closely related work of C. Fefferman and Sánchez-Calle [FSC86].

Tao and Wright [TW03, Section 4] furthered the results of Nagel, Stein, and Wainger and offered a new proof based on methods from ODEs (see Section 9.1 for a detailed discussion of the primary ODE they studied). The second author unified these two approaches to prove more general results in [Str11]. This series of papers can be seen as strengthening and generalizing these theories and casting them in a way which is completely “coordinate free” in the sense that all of our assumptions and estimates are invariant under arbitrary C^2 diffeomorphisms. The most basic version of this scaling perspective can be seen in Example 1.1, as the next example shows.

EXAMPLE 1.2. We take the setting of Example 1.1 with $\mathfrak{M} = \mathbb{R}$, $x_0 = 0$, $X = \delta \frac{\partial}{\partial x}$, for some fixed $\delta > 0$. In this case $\Phi(t) = \delta t$; thus the pullback via Φ is the usual Euclidean dilation of vector fields. We can therefore think of Example 1.1 as a generalization of the usual dilation maps on $\mathbb{R}$.

As described above, the main results of this series have two facets:

- They provide a coordinate system in which given C^1 vector fields have an optimal degree of smoothness.
- They provide a coordinate system in which given vector fields are normalized in a way which is useful for applying techniques from analysis.

These two facets, along with some applications, are described in more detail in Section 7.

Despite the fact that the results in the second paper of this series are sharp in terms of regularity, and the results in this paper are one-derivative off from being optimal, we believe the methods and results of this paper have several advantages over those in the second paper. Some of these advantages are:

- (a) The coordinate system defined in this paper is explicit, while it is only defined implicitly in the second paper.
- (b) The proofs in this paper are simpler. Indeed, the second paper requires all of the results of this paper, plus additional methods from PDEs.
- (c) Despite having a simpler proof, the main results of this paper are still useful in many applications. Indeed, they are stronger, sharper, and more general than the previous works on this subject [NSW85, TW03, Str11] which have had many applications; see Section 7 for further details. However, they are not strong enough to obtain some of the most interesting consequences of the results in the second paper; for example, the results stated in Section 3.2. The PDE methods will also be necessary for future work of the second author in the complex setting; see Section 7.5.
- (d) Because the methods of this paper are based on ODEs, as opposed to the PDEs in the second paper, they are in some ways more robust, and will likely be easier to adapt to other settings. For example, in the third paper of the series we see that these ODE methods can be used to study the real analytic setting.

2 Function Spaces

Before we can state any results, we need to introduce the function spaces we use. We make a distinction between function spaces on subsets of $\mathbb{R}^n$ and function spaces on a C^2 manifold M . On $\mathbb{R}^n$, we have access to the standard coordinate system (and its induced smooth structure) and we can define all of the usual function spaces and their norms in terms of this coordinate system. On M , we do not have access to any such natural coordinates, and it does not make sense to talk about, for example, C^∞ functions on M ; as this would depend on a choice of coordinate system or smooth structure. However, if we are given a finite collection of vector fields on M , it does make sense to talk about functions which are C^∞ with respect to these vector fields, and this is how we shall proceed.

2.1 Function spaces on Euclidean space. Let $\Omega \subset \mathbb{R}^n$ be a bounded, connected, open set (we will almost always be considering the case when Ω is a ball in $\mathbb{R}^n$). We have the following classical Banach spaces of functions on Ω :

$$C(\Omega) = C^0(\Omega) := \{f : \Omega \rightarrow \mathbb{C} \mid f \text{ is continuous and bounded}\},$$

$$\|f\|_{C(\Omega)} = \|f\|_{C^0(\Omega)} := \sup_{x \in \Omega} |f(x)|.$$

For $m \in \mathbb{N}$ (throughout the paper we take the convention $0 \in \mathbb{N}$),

$$C^m(\Omega) := \{f \in C^0(\Omega) \mid \partial_x^\alpha f \in C^0(\Omega), \forall |\alpha| \leq m\}, \quad \|f\|_{C^m(\Omega)} := \sum_{|\alpha| \leq m} \|\partial_x^\alpha f\|_{C^0(\Omega)}.$$

Next we define the classical Lipschitz–Hölder spaces. For $s \in [0, 1]$,

$$\|f\|_{C^{0,s}(\Omega)} := \|f\|_{C^0(\Omega)} + \sup_{\substack{x, y \in \Omega \\ x \neq y}} |x - y|^{-s} |f(x) - f(y)|,$$

$$C^{0,s}(\Omega) := \{f \in C^0(\Omega) : \|f\|_{C^{0,s}(\Omega)} < \infty\}. \quad (2.1)$$

For $m \in \mathbb{N}$, $s \in [0, 1]$,

$$\|f\|_{C^{m,s}(\Omega)} := \sum_{|\alpha| \leq m} \|\partial_x^\alpha f\|_{C^{0,s}(\Omega)}, \quad C^{m,s}(\Omega) := \{f \in C^m(\Omega) : \|f\|_{C^{m,s}(\Omega)} < \infty\}.$$

Next, we turn to the Zygmund–Hölder spaces. Given $h \in \mathbb{R}^n$ define $\Omega_h := \{x \in \mathbb{R}^n : x, x + h, x + 2h \in \Omega\}$. For $s \in (0, 1]$ set

$$\|f\|_{\mathcal{C}^s(\Omega)} := \|f\|_{C^{0,s/2}(\Omega)} + \sup_{\substack{0 \neq h \in \mathbb{R}^n \\ x \in \Omega_h}} |h|^{-s} |f(x + 2h) - 2f(x + h) + f(x)|,$$

$$\mathcal{C}^s(\Omega) := \{f \in C^0(\Omega) : \|f\|_{\mathcal{C}^s(\Omega)} < \infty\}.$$

For $m \in \mathbb{N}$, $s \in (0, 1]$, set

$$\|f\|_{\mathcal{C}^{m+s}(\Omega)} := \sum_{|\alpha| \leq m} \|\partial_x^\alpha f\|_{\mathcal{C}^s(\Omega)}, \quad \mathcal{C}^{m+s}(\Omega) := \{f \in C^m(\Omega) : \|f\|_{\mathcal{C}^{m+s}(\Omega)} < \infty\}.$$

We set,

$$\mathcal{C}^\infty(\Omega) := \bigcap_{s>0} \mathcal{C}^s(\Omega), \quad C^\infty(\Omega) := \bigcap_{m \in \mathbb{N}} C^m(\Omega).$$

When Ω is a ball, $\mathcal{C}^\infty(\Omega) = C^\infty(\Omega)$.

REMARK 2.1. The term $\|f\|_{C^{0,s/2}(\Omega)}$ in the definition of $\|f\|_{\mathcal{C}^s(\Omega)}$ is somewhat unusual, and in the literature is usually replaced by $\|f\|_{C^0(\Omega)}$. As is well-known, if Ω is a bounded Lipschitz domain, these two options yield equivalent² (but not equal) norms—and therefore the space $\mathcal{C}^s(\Omega)$ is the usual Zygmund–Hölder space of order s . However, the constants involved in this equivalence of norms depend on the size of Ω , and the above choice is more convenient for our purposes. For an example of the convenience offered by this choice of norm, see Remark 8.4.

Finally, we turn to spaces of real analytic functions. Given $r > 0$ we define:

$$\|f\|_{C^{\omega,r}(\Omega)} := \sum_{\alpha \in \mathbb{N}^n} \frac{\|\partial_x^\alpha f\|_{C(\Omega)}}{\alpha!} r^{|\alpha|}, \quad C^{\omega,r}(\Omega) := \{f \in C^\infty(\Omega) : \|f\|_{C^{\omega,r}(\Omega)} < \infty\}.$$

We set $C^\omega(\Omega) := \bigcup_{r>0} C^{\omega,r}(\Omega)$. For notational convenience, we set $\mathcal{C}^\omega(\Omega) := C^\omega(\Omega)$.

Throughout the paper, if we say $\|f\|_{C^m(\Omega)} < \infty$, it means that $f \in C^m(\Omega)$, and similarly for any other function space.

For a Banach space V we define the same spaces taking values in V by the obvious modifications and write $C^m(\Omega; V)$, $C^{m,s}(\Omega; V)$, $\mathcal{C}^{m+s}(\Omega; V)$, $C^{\omega,r}(\Omega; V)$, and $C^\omega(\Omega; V)$ to denote these spaces. When we have a vector field X on Ω , we identify X with a function $X : \Omega \rightarrow \mathbb{R}^n$ by writing $X = \sum_{j=1}^n a_j(x) \frac{\partial}{\partial x_j}$ and treating X as the function $X(x) = (a_1(x), \dots, a_n(x))$. Thus, it makes sense to consider norms like $\|X\|_{\mathcal{C}^s(\Omega; \mathbb{R}^n)}$ and $\|X\|_{C^{m,s}(\Omega; \mathbb{R}^n)}$.

2.2 Function spaces on manifolds. Let $X_1, \dots, X_q$ be C^1 vector fields on a connected C^2 manifold M . Define the Carnot–Carathéodory ball associated to $X_1, \dots, X_q$, centered at $x \in M$, of radius $\delta > 0$, by

$$B_X(x, \delta) := \left\{ y \in M \mid \exists \gamma : [0, 1] \rightarrow M, \gamma(0) = x, \gamma(1) = y, \gamma'(t) = \sum_{j=1}^q a_j(t) \delta X_j(\gamma(t)), \right. \\ \left. a_j \in L^\infty([0, 1]), \left\| \sum_{j=1}^q |a_j|^2 \right\|_{L^\infty} < 1 \right\}, \quad (2.2)$$

and for $y \in M$, set

$$\rho(x, y) := \inf\{\delta > 0 : y \in B_X(x, \delta)\}. \quad (2.3)$$

When $X_1, \dots, X_q$ are smooth vector fields on a smooth connected manifold M , if the Lie algebra generated by $X_1, \dots, X_q$ spans the tangent space at every point of

² This equivalence follows easily from [Tri06, Theorem 1.118 (i)]. We will usually use these norms in the case when Ω is a ball in Euclidean space, and is therefore a bounded Lipschitz domain.

M , ρ is a metric on M —sometimes known as a *sub-Riemannian metric* or a *Carnot–Carathéodory metric*. In this case, the metric topology induced by ρ is the same as the topology on M . If the Lie algebra generated by $X_1, \dots, X_q$ does not span the tangent space at some point, then ρ may or may not be a metric: it is possible that $\rho(x, y) = \infty$ for some x, y . If $\rho(x, y) = \infty$, we make the convention that $\rho(x, y)^{-s} = 0$ for $s > 0$ and $\rho(x, y)^0 = 1$. In the nonsmooth setting, we will usually be considering the special case when $X_1, \dots, X_q$ span the tangent space at every point of M , and in this case ρ is a metric, and the metric topology induced by ρ is the same as the topology on M .

We use ordered multi-index notation: X^α . Here, α denotes a list of elements of $\{1, \dots, q\}$ and $|\alpha|$ denotes the length of the list. For example, $X^{(2,1,3,1)} = X_2 X_1 X_3 X_1$ and $|(2, 1, 3, 1)| = 4$.

Associated to the vector fields $X_1, \dots, X_q$, we have the following Banach spaces of functions on M .

$$C(M) = C_X^0(M) := \{f : M \rightarrow \mathbb{C} \mid f \text{ is continuous and bounded}\},$$

$$\|f\|_{C(M)} = \|f\|_{C_X^0(M)} := \sup_{x \in M} |f(x)|.$$

For $m \in \mathbb{N}$, we define

$$C_X^m(M) := \{f \in C(M) \mid X^\alpha f \text{ exists and } X^\alpha f \in C(M), \forall |\alpha| \leq m\},$$

$$\|f\|_{C_X^m(M)} := \sum_{|\alpha| \leq m} \|X^\alpha f\|_{C(M)}.$$

For $s \in [0, 1]$, we define the Lipschitz–Hölder space associated to X by

$$\|f\|_{C_X^{0,s}(M)} := \|f\|_{C(M)} + \sup_{\substack{x, y \in M \\ x \neq y}} \rho(x, y)^{-s} |f(x) - f(y)|,$$

$$C_X^{0,s}(M) := \{f \in C(M) : \|f\|_{C_X^{0,s}(M)} < \infty\}.$$

For $m \in \mathbb{N}$ and $s \in [0, 1]$, set

$$\|f\|_{C_X^{m,s}(M)} := \sum_{|\alpha| \leq m} \|X^\alpha f\|_{C_X^{0,s}(M)}, \quad C_X^{m,s}(M) := \{f \in C_X^m(M) : \|f\|_{C_X^{m,s}(M)} < \infty\}.$$

We turn to the Zygmund–Hölder spaces. For this, we use the Hölder spaces $C^{0,s}([a, b])$ for a closed interval $[a, b] \subset \mathbb{R}$; $\|\cdot\|_{C^{0,s}([a, b])}$ is defined via the formula (2.1). Given $h > 0$, $s \in (0, 1)$ define

$$\mathcal{P}_{X,s}^M(h) := \left\{ \gamma : [0, 2h] \rightarrow M \mid \gamma'(t) = \sum_{j=1}^q d_j(t) X_j(\gamma(t)), d_j \in C^{0,s}([0, 2h]), \right.$$

$$\left. \sum_{j=1}^q \|d_j\|_{C^{0,s}([0, 2h])}^2 < 1 \right\}.$$

For $s \in (0, 1]$ set

$$\|f\|_{\mathcal{C}_X^s(M)} := \|f\|_{C_X^{0,s/2}(M)} + \sup_{\substack{h>0 \\ \gamma \in \mathcal{P}_{X,s/2}^M(h)}} h^{-s} |f(\gamma(2h)) - 2f(\gamma(h)) + f(\gamma(0))|,$$

and for $m \in \mathbb{N}$,

$$\|f\|_{\mathcal{C}_X^{m+s}(M)} := \sum_{|\alpha| \leq m} \|X^\alpha f\|_{\mathcal{C}_X^s(M)},$$

and we set

$$\mathcal{C}_X^{s+m}(M) := \{f \in C_X^m(M) : \|f\|_{\mathcal{C}_X^{m+s}(M)} < \infty\}.$$

Set

$$\mathcal{C}_X^\infty(M) := \bigcap_{s>0} \mathcal{C}_X^s(M) \text{ and } C_X^\infty(M) := \bigcap_{m \in \mathbb{N}} C_X^m(M).$$

It is a consequence of Lemma 8.1 that $\mathcal{C}_X^\infty(M) = C_X^\infty(M)$; indeed, $\mathcal{C}_X^\infty(M) \subseteq C_X^\infty(M)$ is clear while the reverse containment follows from Lemma 8.1.

We introduce the following counter-intuitive, but convenient, definitions.

DEFINITION 2.2. For $m < 0$, $s \in [0, 1]$, we define $C_X^{m,s}(M) := C(M)$ with equality of norms. For $s \in (-1, 0]$, we define $\mathcal{C}_X^s(M) := C_X^{0,(s+1)/2}(M)$, with equality of norms. For $s \in (-\infty, -1]$, we define $\mathcal{C}_X^s(M) := C(M)$ with equality of norms.

Finally, for $r > 0$ we introduce a space of functions which are “real analytic with respect to X ”.

$$\|f\|_{C_X^{\omega,r}(M)} := \sum_{m=0}^{\infty} \frac{r^m}{m!} \sum_{|\alpha|=m} \|X^\alpha f\|_{C(M)}, \quad C_X^{\omega,r}(M) := \{f \in C_X^\infty(M) : \|f\|_{C_X^{\omega,r}(M)} < \infty\}.$$

This definition was introduced in greater generality by Nelson [Nel59]. We set $C_X^\omega(M) := \bigcup_{r>0} C_X^{\omega,r}(M)$, and for notational convenience set $\mathcal{C}_X^\omega(M) := C_X^\omega(M)$. We refer the reader to the third paper in the series for a more detailed discussion of the spaces $C^{\omega,r}(\Omega)$ and $C_X^{\omega,r}(M)$.

Importantly, all of the above spaces are invariant under diffeomorphisms. In fact, we have the following result.

PROPOSITION 2.3. Let N be another C^2 manifold, let $\Phi : M \rightarrow N$ be a C^2 diffeomorphism, and let Φ_*X denote the list of vector fields $\Phi_*X_1, \dots, \Phi_*X_q$. Then the map $f \mapsto f \circ \Phi$ is an isometric isomorphism between the following Banach spaces: $C_{\Phi_*X}^m(N) \rightarrow C_X^m(M)$, $C_{\Phi_*X}^{m,s}(N) \rightarrow C_X^{m,s}(M)$, $\mathcal{C}_{\Phi_*X}^s(N) \rightarrow \mathcal{C}_X^s(M)$, and $C_{\Phi_*X}^{\omega,r}(N) \rightarrow C_X^{\omega,r}(M)$.

Proof. This is immediate from the definitions. □

REMARK 2.4. Some of the above definitions deserve some additional remarks.

- In (2.2), $\gamma'(t)$ is defined as follows. In the case that M is an open subset $\Omega \subseteq \mathbb{R}^n$ and $\gamma : [a, b] \rightarrow \Omega$, $\gamma'(t) = \sum_{j=1}^q a_j(t)X_j(\gamma(t))$ is defined to mean $\gamma(t) = \gamma(a) + \int_a^t \sum_j a_j(s)X_j(\gamma(s)) ds$; note that this definition is local in t . For an abstract C^2 manifold M , this is interpreted locally. I.e., if $\gamma : [a, b] \rightarrow M$, we say $\gamma'(t) = \sum_{j=1}^q a_j(t)X_j(\gamma(t))$ if $\forall t_0 \in [a, b]$, there is an open neighborhood N of $\gamma(t_0)$ and a C^2 diffeomorphism $\Psi : N \rightarrow \Omega$, where $\Omega \subseteq \mathbb{R}^n$ is open, such that $(\Psi \circ \gamma)'(t) = \sum_{j=1}^q a_j(t)(\Psi_* X_j)(\Psi \circ \gamma(t))$ for t near t_0 ($t \in [a, b]$).
- When we write Vf for a C^1 vector field V and $f : M \rightarrow \mathbb{R}$, we define this as $Vf(x) := \frac{d}{dt}\big|_{t=0} f(e^{tV}x)$. When we say Vf exists, it means that this derivative exists in the classical sense, $\forall x$. If we have several C^1 vector fields $V_1, V_2, \dots, V_L$, we define $V_1 V_2 \cdots V_L f := V_1(V_2(\cdots V_L(f)))$ and to say that this exists means that at each stage the derivatives exist.

2.2.1 Beyond manifolds. For certain subsets of M which are not themselves manifolds, we can still define the above norms. Indeed, let $X_1, \dots, X_q$ be C^1 vector fields on a C^2 manifold M and fix $\xi > 0$. In this setting, $B_X(x_0, \xi)$ might not be a manifold (though it sometimes is—see Proposition 3.1). $B_X(x_0, \xi)$ is a metric space, with the metric ρ . For a function $f : B_X(x_0, \xi) \rightarrow \mathbb{C}$ and $x \in B_X(x_0, \xi)$, it makes sense to consider $X_j f(x) := \frac{d}{dt}\big|_{t=0} f(e^{tX_j}x)$. Using this, we can define the spaces $C_X^{m,s}(B_X(x_0, \xi))$, $\mathcal{C}_X^s(B_X(x_0, \xi))$, and $C_X^{\omega,r}(B_X(x_0, \xi))$ and their corresponding norms, with the same formulas as above.

3 Overview of the Series

In this section, we present the main results of this three part series of papers; though we will offer a more detailed presentation of these results in the later papers. We separate the results into two parts: the qualitative results (i.e., (i) and (ii) from the introduction) and the quantitative results (i.e., (iii)). The quantitative results are the most useful for applications, and the qualitative results are simple consequences of the quantitative ones. The proofs will not be completed until the later papers—though in this paper we prove a slightly weaker version of the quantitative results (see Section 4). We begin by stating the qualitative results, as they are easier to understand.

3.1 Qualitative results. Let $X_1, \dots, X_q$ be C^1 vector fields on a C^2 manifold $\mathfrak{M}$. For $x, y \in \mathfrak{M}$, define $\rho(x, y)$ as in (2.3). Fix $x_0 \in \mathfrak{M}$ and let $Z := \{y \in \mathfrak{M} : \rho(x_0, y) < \infty\}$. ρ is a metric on Z , and we give Z the topology induced by ρ (this is finer³ than the topology as a subspace of $\mathfrak{M}$, and may be strictly finer). Let $M \subseteq Z$ be a connected open subset of Z containing x_0 . We give M the topology of a subspace of Z . We begin with a classical result to set the stage.

³ See Lemma A.1 for a proof that this topology is finer than the subspace topology.

PROPOSITION 3.1. Suppose $[X_i, X_j] = \sum_{k=1}^q c_{i,j}^k X_k$, where $c_{i,j}^k : M \rightarrow \mathbb{R}$ are locally bounded. Then, there is a C^2 manifold structure on M (compatible with its topology) such that:

- The inclusion $M \hookrightarrow \mathfrak{M}$ is a C^2 injective immersion.
- $X_1, \dots, X_q$ are C^1 vector fields tangent to M .
- $X_1, \dots, X_q$ span the tangent space at every point of M .

Furthermore, this C^2 structure is unique in the sense that if M is given another C^2 structure (compatible with its topology) such that the inclusion map $M \hookrightarrow \mathfrak{M}$ is a C^2 injective immersion, then the identity map $M \rightarrow M$ is a C^2 diffeomorphism between these two structures.

For a proof of Proposition 3.1 see Appendix A. Henceforth, we assume the conditions of Proposition 3.1 so that M is a C^2 manifold and $X_1, \dots, X_q$ are C^1 vector fields on M which span the tangent space at every point. We write $n = \dim \text{span}\{X_1(x_0), \dots, X_q(x_0)\}$, so that $\dim M = n$.

REMARK 3.2. If $X_1(x_0), \dots, X_q(x_0)$ span $T_{x_0}\mathfrak{M}$, then M is an open submanifold of $\mathfrak{M}$. If $X_1, \dots, X_q$ span the tangent space at every point of $\mathfrak{M}$ and $\mathfrak{M}$ is connected, one may take $M = \mathfrak{M}$.

Theorem 3.3 (The Local Theorem). For $s \in (1, \infty] \cup \{\omega\}$, the following three conditions are equivalent:

- (i) There is an open neighborhood $V \subseteq M$ of x_0 and a C^2 diffeomorphism $\Phi : U \rightarrow V$ where $U \subseteq \mathbb{R}^n$ is open, such that $\Phi^*X_1, \dots, \Phi^*X_q \in \mathcal{C}^{s+1}(U; \mathbb{R}^n)$.
- (ii) Re-order the vector fields so that $X_1(x_0), \dots, X_n(x_0)$ are linearly independent. There is an open neighborhood $V \subseteq M$ of x_0 such that:
 - $[X_i, X_j] = \sum_{k=1}^n \hat{c}_{i,j}^k X_k$, $1 \leq i, j \leq n$, where $\hat{c}_{i,j}^k \in \mathcal{C}_X^s(V)$.
 - For $n+1 \leq j \leq q$, $X_j = \sum_{k=1}^n b_j^k X_k$, where $b_j^k \in \mathcal{C}_X^{s+1}(V)$.
- (iii) There exists an open neighborhood $V \subseteq M$ of x_0 such that $[X_i, X_j] = \sum_{k=1}^q c_{i,j}^k X_k$, $1 \leq i, j \leq q$, where $c_{i,j}^k \in \mathcal{C}_X^s(V)$.

REMARK 3.4. (ii) and (iii) of Theorem 3.3 are similar but have slightly different advantages. In (ii), because $X_1, \dots, X_n$ form a basis for the tangent space of M near x_0 , the functions $\hat{c}_{i,j}^k$ and b_j^k are uniquely determined (so long as V is chosen sufficiently small), and one can directly check to see if (ii) holds by computing these functions.⁴ If $q > n$, $X_1, \dots, X_q$ are linearly dependent, so the $c_{i,j}^k$ in (iii) are not unique—and (iii) only asks that there exists a choice of $c_{i,j}^k$ satisfying the conditions in (iii). Despite this lack of uniqueness, (iii) is the setting which usually arises in applications.

⁴ The computation can be done in any coordinate system, as the conditions are invariant under a change of coordinate system—see Proposition 2.3.

REMARK 3.5. Theorem 3.3 is stated for $s \in (1, \infty]$. It is reasonable to expect the same result for $s \in (0, \infty]$, however our proof runs into some technical issues when $s \in (0, 1]$. We refer the reader to the second paper for a further discussion of this. A similar remark holds for Theorem 3.6, below.

Theorem 3.6 (The Global Theorem). *For $s \in (1, \infty]$, the following three conditions are equivalent:*

- (i) *There exists a $\mathcal{C}^{s+2}$ atlas on M , compatible with its C^2 structure, such that $X_1, \dots, X_q$ are $\mathcal{C}^{s+1}$ with respect to this atlas.*
- (ii) *For each $x_0 \in M$, any of the three equivalent conditions (i), (ii), or (iii) from Theorem 3.3 holds for this choice of x_0 .*
- (iii) *$[X_i, X_j] = \sum_{k=1}^q c_{i,j}^k X_k$, $1 \leq i, j \leq q$, where $\forall x_0 \in M$, $\exists V \subseteq M$ open with $x_0 \in V$ such that $c_{i,j}^k|_V \in \mathcal{C}_X^s(V)$, $1 \leq i, j, k \leq q$.*

Furthermore, under these conditions, the $\mathcal{C}^{s+2}$ manifold structure on M induced by the atlas in (i) is unique, in the sense that if there is another $\mathcal{C}^{s+2}$ atlas on M , compatible with its C^2 structure, and such that $X_1, \dots, X_q$ are $\mathcal{C}^{s+1}$ with respect to this second atlas, then the identity map $M \rightarrow M$ is a $\mathcal{C}^{s+2}$ diffeomorphism between these two $\mathcal{C}^{s+2}$ manifold structures on M .

Also, the following two conditions are equivalent:

- (a) *There is a real analytic atlas on M , compatible with its C^2 structure, such that $X_1, \dots, X_q$ are real analytic with respect to this atlas.*
- (b) *For each $x_0 \in M$, any of the three equivalent conditions (i), (ii), or (iii) from Theorem 3.3 hold for this choice of x_0 (with $s = \omega$).*

Furthermore, under these conditions, the real analytic manifold structure on M induced by the atlas in (a) is unique, in the sense that if there is another real analytic atlas on M , compatible with its C^2 structure and such that $X_1, \dots, X_q$ are real analytic with respect to this second atlas, then the identity map $M \rightarrow M$ is a real analytic diffeomorphism between these two real analytic structures on M .

3.2 Quantitative results. Theorem 3.3 gives necessary and sufficient conditions for a certain type of coordinate chart to exist. For applications in analysis, it is essential to have quantitative control of this coordinate chart. In the second part to this series, these quantitative charts are studied in the setting of Zygmund spaces, while in the third part they are studied in the real analytic setting. In this section, we present the results on Zygmund spaces, and refer the reader to the third paper for the corresponding real analytic results.

Because we need to keep track of what each constant depends on for applications in analysis (see Section 7), the statements of the results in this section, later in the paper, and in the subsequent papers in this series, are quite technical. To help simplify matters, we define various notions of “admissible constants”. These will be constants that can only depend on certain parameters. While these definitions are somewhat unwieldy, they greatly simplify the statements of the results in the rest

of this series. In each instance, it will be clear what notion of admissible constants we are using.

First we need some new notation. $B^n(\eta)$ denotes the Euclidean ball of radius $\eta > 0$ centered at $0 \in \mathbb{R}^n$. Let $X_1, \dots, X_q$ be C^1 vector fields on a C^2 manifold $\mathfrak{M}$.

DEFINITION 3.7. For $x_0 \in \mathfrak{M}$, $\eta > 0$, and $U \subseteq \mathfrak{M}$, we say the list $X = X_1, \dots, X_q$ satisfies $\mathcal{C}(x_0, \eta, U)$ if for every $a \in B^q(\eta)$ the expression

$$e^{a_1 X_1 + \dots + a_q X_q} x_0$$

exists in U . More precisely, consider the differential equation

$$\frac{\partial}{\partial r} E(r) = a_1 X_1(E(r)) + \dots + a_q X_q(E(r)), \quad E(0) = x_0.$$

We assume that a solution to this differential equation exists up to $r = 1$, $E : [0, 1] \rightarrow U$. We have $E(r) = e^{ra_1 X_1 + \dots + ra_q X_q} x_0$.

For $1 \leq n \leq q$, we let

$$\begin{aligned} \mathcal{I}(n, q) &:= \{(i_1, i_2, \dots, i_n) : i_j \in \{1, \dots, q\}\}, \\ \mathcal{I}_0(n, q) &:= \{i \in \mathcal{I}(n, q) : 1 \leq i_1 < i_2 < \dots < i_n \leq q\}. \end{aligned}$$

For $J = (j_1, \dots, j_n) \in \mathcal{I}(n, q)$ we write X_J for the list of vector fields $X_{j_1}, \dots, X_{j_n}$. We write $\bigwedge X_J = X_{j_1} \wedge X_{j_2} \wedge \dots \wedge X_{j_n}$.

Fix $x_0 \in \mathfrak{M}$, let $n = \dim \text{span}\{X_1(x_0), \dots, X_q(x_0)\}$. Fix $\xi, \zeta \in (0, 1]$. We assume that on $B_X(x_0, \xi)$, the X_j 's satisfy

$$[X_j, X_k] = \sum_{l=1}^q c_{j,k}^l X_l, \quad c_{j,k}^l \in C(B_X(x_0, \xi)),$$

where $B_X(x_0, \xi)$ is given the metric topology induced by ρ from (2.3). Proposition 3.1 applies to show that $B_X(x_0, \xi)$ is an n -dimensional, C^2 , injectively immersed submanifold of $\mathfrak{M}$. $X_1, \dots, X_q$ are C^1 vector fields on $B_X(x_0, \xi)$ and span the tangent space at every point. Henceforth, we treat $X_1, \dots, X_q$ as vector fields on $B_X(x_0, \xi)$.

Let $J_0 \in \mathcal{I}(n, q)$ be such that $\bigwedge X_{J_0}(x_0) \neq 0$ and moreover

$$\max_{J \in \mathcal{I}(n, q)} \left| \frac{\bigwedge X_J(x_0)}{\bigwedge X_{J_0}(x_0)} \right| \leq \zeta^{-1}, \quad (3.1)$$

see Section 5 for the definition of this quotient. Note that such a $J_0 \in \mathcal{I}(n, q)$ always exists—indeed, we may choose J_0 so that the left hand side of (3.1) equals 1. Without loss of generality, reorder the vector fields so that $J_0 = (1, \dots, n)$.

- Let $\eta > 0$ be such that X_{J_0} satisfies $\mathcal{C}(x_0, \eta, \mathfrak{M})$.
- Let $\delta_0 > 0$ be such that for $\delta \in (0, \delta_0]$ the following holds: if $z \in B_{X_{J_0}}(x_0, \xi)$ is such that X_{J_0} satisfies $\mathcal{C}(z, \delta, B_{X_{J_0}}(x_0, \xi))$ and if $t \in B^n(\delta)$ is such that $e^{t_1 X_1 + \dots + t_n X_n} z = z$ and if $X_1(z), \dots, X_n(z)$ are linearly independent, then $t = 0$.

REMARK 3.8. Using that the vector fields $X_1, \dots, X_n$ are C^1 , it follows that there exist η and δ_0 as above (which are small depending on, among other things, the C^1 norms of $X_1, \dots, X_n$ in a fixed coordinate system); see Proposition 4.14. However, it is possible that the C^1 norms of $X_1, \dots, X_q$ can be very large while η and δ_0 are not small. Furthermore, the quantities η and δ_0 are invariant under C^2 diffeomorphisms, while the C^1 norms of $X_1, \dots, X_n$ depend on the choice of coordinate system. Thus, we present our results in terms of η and δ_0 .

REMARK 3.9. For a more detailed discussion of η and δ_0 see Section 4.1.

Fix $s_0 > 1$.

DEFINITION 3.10. For $s \geq s_0$ if we say C is an $\{s\}$ -admissible constant, it means that we assume $c_{j,k}^l \in \mathcal{C}_{X_{J_0}}^s(B_{X_{J_0}}(x_0, \xi))$ for $1 \leq j, k, l \leq q$. C is then allowed to depend on s, s_0 , lower bounds > 0 for ζ, ξ, η , and δ_0 , and upper bounds for q and $\|c_{j,k}^l\|_{\mathcal{C}_{X_{J_0}}^s(B_{X_{J_0}}(x_0, \xi))}$, $1 \leq j, k, l \leq q$. We write $A \lesssim_{\{s\}} B$ for $A \leq CB$ where C is a positive $\{s\}$ -admissible constant. We write $A \approx_{\{s\}} B$ for $A \lesssim_{\{s\}} B$ and $B \lesssim_{\{s\}} A$.

Theorem 3.11 (The Quantitative Theorem). Suppose $c_{i,j}^k \in \mathcal{C}_X^{s_0}(B_{X_{J_0}}(x_0, \xi))$, $1 \leq i, j, k \leq q$. Then, there exists a map $\Phi : B^n(1) \rightarrow B_{X_{J_0}}(x_0, \xi)$ and $\{s_0\}$ -admissible constants $\xi_1, \xi_2 > 0$ such that the following hold:

- (i) $\Phi(B^n(1)) \subseteq B_X(x_0, \xi)$ is an open subset of the C^2 manifold $B_X(x_0, \xi)$.
- (ii) $\Phi : B^n(1) \rightarrow \Phi(B^n(1))$ is a C^2 diffeomorphism.
- (iii) $B_X(x_0, \xi_2) \subseteq B_{X_{J_0}}(x_0, \xi_1) \subseteq \Phi(B^n(1)) \subseteq B_X(x_0, \xi)$.

Let $Y_j = \Phi^* X_j$. There exists an $\{s_0\}$ -admissible constant $K \approx_{\{s_0\}} 1$ and a matrix $A \in \mathcal{C}^{s_0}(B^n(1); \mathbb{M}^{n \times n})$ such that:⁵

- (iv) $Y_{J_0} = K(I + A)\nabla$, where ∇ denotes the gradient in $\mathbb{R}^n$ (thought of as a column vector) and we are identifying Y_{J_0} with the column vector of vector fields $[Y_1, Y_2, \dots, Y_n]^\top$.
- (v) $A(0) = 0$, $\sup_{t \in B^n(1)} \|A(t)\|_{\mathbb{M}^{n \times n}} \leq \frac{1}{2}$.
- (vi) For all $s \geq s_0$, $1 \leq j \leq q$, $\|Y_j\|_{\mathcal{C}^{s+1}(B^n(1); \mathbb{R}^n)} \lesssim_{\{s\}} 1$.

REMARK 3.12. In the second paper, we discuss further details of the map Φ from Theorem 3.11. For example, we describe how to understand $\Phi^* \nu$ where ν is a density on $B_X(x_0, \xi)$.

3.2.1 Diffeomorphism invariance. The results in this series are invariant under arbitrary C^2 diffeomorphisms. In light of Proposition 2.3 this is obvious for the qualitative results (Theorems 3.3 and 3.6). It is true for the quantitative results as well (e.g., Theorem 3.11).

Indeed, let $X_1, \dots, X_q$ be C^1 vector fields on a C^2 manifold $\mathfrak{M}$, as in Theorem 3.11, and fix $x_0 \in \mathfrak{M}$. Let $\Psi : \mathfrak{M} \rightarrow \mathfrak{N}$ be a C^2 diffeomorphism. Then, $X_1, \dots, X_q$

⁵ Here, and in the rest of the paper, $\mathbb{M}^{n \times n}$ denotes the space of $n \times n$ real matrices endowed with the usual operator norm of a matrix.

satisfy the conditions of Theorem 3.11 at the point x_0 if and only if $\Psi_*X_1, \dots, \Psi_*X_q$ satisfy the conditions at $\Psi(x_0)$. Moreover, $\{s\}$ -admissible constants as defined in terms of $X_1, \dots, X_q$ are the same as $\{s\}$ -admissible constants when defined in terms of $\Psi_*X_1, \dots, \Psi_*X_q$. Finally, if Φ is the map guaranteed by Theorem 3.11 when applied to $X_1, \dots, X_q$, then $\Psi \circ \Phi$ is the map guaranteed by Theorem 3.11 when applied to $\Psi_*X_1, \dots, \Psi_*X_q$ (as can be seen by tracing through the proof). The same remarks hold for Theorem 4.7, below.

4 Main Results of this Paper

We now turn to the results of this paper, which amount to a slightly weaker version of Theorem 3.11. We take the same setup as Theorem 3.11; so that we have $X_1, \dots, X_q$, C^1 vector fields on a C^2 manifold $\mathfrak{M}$. Fix $x_0 \in M$ and set $n = \dim \text{span}\{X_1(x_0), \dots, X_q(x_0)\}$. As before, we assume that on $B_X(x_0, \xi)$, the X_j 's satisfy

$$[X_j, X_k] = \sum_{l=1}^q c_{j,k}^l X_l, \quad c_{j,k}^l \in C(B_X(x_0, \xi)),$$

where $B_X(x_0, \xi)$ is given the metric topology induced by ρ from (2.3). Proposition 3.1 applies to show that $B_X(x_0, \xi)$ is an n -dimensional, C^2 , injectively immersed submanifold of $\mathfrak{M}$. $X_1, \dots, X_q$ are C^2 vector fields on $B_X(x_0, \xi)$ and span the tangent space at every point. Henceforth, we treat $X_1, \dots, X_q$ as vector fields on $B_X(x_0, \xi)$. Let $J_0 \in \mathcal{I}(n, q)$ be such that $\bigwedge X_{J_0}(x_0) \neq 0$ and moreover

$$\max_{J \in \mathcal{I}(n, q)} \left| \frac{\bigwedge X_J(x_0)}{\bigwedge X_{J_0}(x_0)} \right| \leq \zeta^{-1},$$

see Section 5 for the definition of this quotient.⁶ Without loss of generality, reorder the vector fields so that $J_0 = (1, \dots, n)$. Let $\eta, \delta_0 > 0$ be as in Section 3.2.

DEFINITION 4.1. We say C is a 0-admissible constant if C can be chosen to depend only on upper bounds for q , ζ^{-1} , ξ^{-1} , and $\|c_{j,k}^l\|_{C(B_{X_{J_0}}(x_0, \xi))}$, $1 \leq j, k, l \leq q$.

DEFINITION 4.2. If we say C is a 1-admissible constant, it means that we assume $c_{j,k}^l \in C_X^1(B_{X_{J_0}}(x_0, \xi))$ for $1 \leq j, k \leq n$, $1 \leq l \leq q$. C is then allowed to depend on anything a 0-admissible constant can depend on, lower bounds > 0 for η and δ_0 , and upper bounds for $\|c_{j,k}^l\|_{C_X^1(B_{X_{J_0}}(x_0, \xi))}$, $1 \leq j, k \leq n$, $1 \leq l \leq q$.

DEFINITION 4.3. For $m_1, m_2 \in \mathbb{Z}$ and $s \in [0, 1]$ if we say C is an $\langle m_1, m_2, s \rangle$ -admissible constant, it means that we assume:

⁶ One may always choose J_0 so that $\zeta = 1$. However, the flexibility to take $\zeta < 1$ is essential for some applications. It will prove to be particularly important when we turn to analogous results in the complex setting in a future paper.

- $c_{j,k}^l \in C_{X_{J_0}}^{m_1,s}(B_{X_{J_0}}(x_0, \xi)), 1 \leq j, k \leq n, 1 \leq l \leq q.$
- $c_{j,k}^l \in C_{X_{J_0}}^{m_2,s}(B_{X_{J_0}}(x_0, \xi)), 1 \leq j, k, l \leq q.$

C can then be chosen to depend only on upper bounds for $m_1, m_2, q, \zeta^{-1}, \xi^{-1}, \|c_{j,k}^l\|_{C_{X_{J_0}}^{m_1,s}(B_{X_{J_0}}(x_0, \xi))}, 1 \leq j, k \leq n, 1 \leq l \leq q$, and $\|c_{j,k}^l\|_{C_{X_{J_0}}^{m_2,s}(B_{X_{J_0}}(x_0, \xi))}, 1 \leq j, k, l \leq q.$

DEFINITION 4.4. For $s_1, s_2 \in \mathbb{R}$ if we say C is an $\{s_1, s_2\}$ -admissible constant, it means that we assume:

- $c_{j,k}^l \in \mathcal{C}_{X_{J_0}}^{s_1}(B_{X_{J_0}}(x_0, \xi)), 1 \leq j, k \leq n, 1 \leq l \leq q.$
- $c_{j,k}^l \in \mathcal{C}_{X_{J_0}}^{s_2}(B_{X_{J_0}}(x_0, \xi)), 1 \leq j, k, l \leq q.$

C can then be chosen to depend only on s_1, s_2 and upper bounds for $q, \zeta^{-1}, \eta^{-1}, \xi^{-1}, \|c_{j,k}^l\|_{\mathcal{C}_{X_{J_0}}^{s_1}(B_{X_{J_0}}(x_0, \xi))}, 1 \leq j, k \leq n, 1 \leq l \leq q$, and $\|c_{j,k}^l\|_{\mathcal{C}_{X_{J_0}}^{s_2}(B_{X_{J_0}}(x_0, \xi))}, 1 \leq j, k, l \leq q.$

REMARK 4.5. 0 and 1-admissible constants are the most basic type of admissible constants, and nearly all of our estimates depend on those quantities used in 0-admissible constants, while many depend on the stronger 1-admissible constants. Admissible constants using the braces $\langle \cdot \rangle$ are used when working with estimates relating to Hölder norms, while those using $\{ \cdot \}$ are used for estimates relating to Zygmund norms. In Section 6, we introduce a density ν and admissible constants that take into account this density. To indicate this, we will decorate the notions of admissible constants by writing, e.g., $\langle m_1, m_2, s; \nu \rangle$ -admissible constants and $\{s_1, s_1; \nu\}$ -admissible constants. Finally, in Section 8.1 we will prove some technical results for vector fields which are defined on Euclidean space. To indicate the corresponding admissible constants, we will use notation like $\langle m_1, s; E \rangle$ and $\{s; E\}$, where E stands for “Euclidean”.

REMARK 4.6. In the various definitions of admissible constants in this section, we treat $c_{j,k}^l$ differently depending on whether $1 \leq j, k \leq n$ or $1 \leq j, k \leq q$. This is likely an artifact of the proof. Indeed, this lack of symmetry disappears when we move to the sharp results in the second paper in the series; see Theorem 3.11.

We write $A \lesssim_0 B$ for $A \leq CB$ where C is a positive 0-admissible constant. We write $A \approx_0 B$ for $A \lesssim_0 B$ and $B \lesssim_0 A$. We similarly define $\lesssim_1, \approx_1, \lesssim_{\langle m_1, m_2, s \rangle}, \approx_{\langle m_1, m_2, s \rangle}, \lesssim_{\{s_1, s_2\}},$ and $\approx_{\{s_1, s_2\}}.$

Because X_{J_0} satisfies $\mathcal{C}(x_0, \eta, \mathfrak{M})$, by hypothesis, we may define the map, for $t \in B^n(\eta)$,

$$\Phi(t) := e^{t_1 X_1 + \dots + t_n X_n} x_0. \quad (4.1)$$

Let $\eta_0 := \min\{\eta, \xi\}$ so that $\Phi : B^n(\eta_0) \rightarrow B_{X_{J_0}}(x_0, \xi)$. Note that, a priori, Φ is C^1 , since $X_1, \dots, X_n$ are C^1 .

Theorem 4.7. *There exists a 0-admissible constant $\chi \in (0, \xi]$ such that:*

- (a) $\forall y \in B_{X_{J_0}}(x_0, \chi), \bigwedge X_{J_0}(y) \neq 0$.
- (b) $\forall y \in B_{X_{J_0}}(x_0, \chi),$

$$\max_{J \in \mathcal{I}(n, q)} \left| \frac{\bigwedge X_J(y)}{\bigwedge X_{J_0}(y)} \right| \approx_0 1.$$

- (c) $\forall \chi' \in (0, \chi], B_{X_{J_0}}(x_0, \chi')$ is an open subset of $B_X(x_0, \xi)$ and is therefore a submanifold.

For the rest of the theorem, we assume $c_{j,k}^l \in C_{X_{J_0}}^1(B_{X_{J_0}}(x_0, \xi))$ for $1 \leq j, k \leq n$, $1 \leq l \leq q$. There exist 1-admissible constants $\eta_1, \xi_1, \xi_2 > 0$ such that:

- (d) $\Phi(B^n(\eta_1))$ is an open subset of $B_{X_{J_0}}(x_0, \chi)$, and is therefore a submanifold of $B_X(x_0, \xi)$.
- (e) $\Phi : B^n(\eta_1) \rightarrow \Phi(B^n(\eta_1))$ is a C^2 diffeomorphism.
- (f) $B_X(x_0, \xi_2) \subseteq B_{X_{J_0}}(x_0, \xi_1) \subseteq \Phi(B^n(\eta_1)) \subseteq B_{X_{J_0}}(x_0, \chi) \subseteq B_X(x_0, \xi)$.

Let $Y_j = \Phi^* X_j$ and write $Y_{J_0} = (I + A)\nabla$, where Y_{J_0} denotes the column vector of vector fields $Y_{J_0} = [Y_1, Y_2, \dots, Y_n]^\top$, ∇ denotes the gradient in $\mathbb{R}^n$ thought of as a column vector, and $A \in C(B^n(\eta_1); \mathbb{M}^{n \times n})$.

- (g) $A(0) = 0$ and $\sup_{t \in B^n(\eta_1)} \|A(t)\|_{\mathbb{M}^{n \times n}} \leq \frac{1}{2}$.
- (h) We have the following regularity on Y_j , $1 \leq j \leq q$:
 - $\|Y_j\|_{C^{m,s}(B^n(\eta_1); \mathbb{R}^n)} \lesssim_{\langle m, m-1, s \rangle} 1$, for $m \in \mathbb{N}$, $s \in [0, 1]$.
 - $\|Y_j\|_{\mathcal{C}^s(B^n(\eta_1); \mathbb{R}^n)} \lesssim_{\{s, s-1\}} 1$, for $s > 0$.
- (i) There exist $b_k^l \in C^1(B^n(\eta_1))$, $n+1 \leq k \leq q$, $1 \leq l \leq n$, such that $Y_k = \sum_{l=1}^n b_k^l Y_l$ and

$$\begin{aligned} \|b_k^l\|_{C^{m,s}(B^n(\eta_1))} &\lesssim_{\langle m-1, m-1, s \rangle} 1, \quad m \in \mathbb{N}, s \in [0, 1], \\ \|b_k^l\|_{\mathcal{C}^s(B^n(\eta_1))} &\lesssim_{\{s-1, s-1\}} 1, \quad s > 0. \end{aligned}$$

- (j) For $1 \leq j, k \leq n$, $[Y_j, Y_k] = \sum_{l=1}^n \tilde{c}_{j,k}^l Y_l$, where

$$\begin{aligned} \|\tilde{c}_{j,k}^l\|_{C^{m,s}(B^n(\eta_1))} &\lesssim_{\langle m, m-1, s \rangle} 1, \quad m \in \mathbb{N}, s \in [0, 1], \\ \|\tilde{c}_{j,k}^l\|_{\mathcal{C}^s(B^n(\eta_1))} &\lesssim_{\{s, s-1\}} 1, \quad s > 0. \end{aligned}$$

- (k) We have the following equivalence of norms, for $f \in C(B^n(\eta_1))$,
 - $\|f\|_{C^{m,s}(B^n(\eta_1))} \approx_{\langle m-1, m-2, s \rangle} \|f\|_{C_{Y_{J_0}}^{m,s}(B^n(\eta_1))} \approx_{\langle m-1, m-2, s \rangle} \|f\|_{C_Y^{m,s}(B^n(\eta_1))}$, for $m \in \mathbb{N}$, $s \in [0, 1]$.
 - $\|f\|_{\mathcal{C}^s(B^n(\eta_1))} \approx_{\{s-1, s-2\}} \|f\|_{\mathcal{C}_{Y_{J_0}}^s(B^n(\eta_1))} \approx_{\{s-1, s-2\}} \|f\|_{\mathcal{C}_Y^s(B^n(\eta_1))}$, for $s > 2$.
- (l) We have, for $f \in C(B_{X_{J_0}}(x_0, \chi))$,
 - $\|f \circ \Phi\|_{C^{m,s}(B^n(\eta_1))} \lesssim_{\langle m-1, m-2, s \rangle} \|f\|_{C_{X_{J_0}}^{m,s}(B_{X_{J_0}}(x_0, \chi))}$, $m \in \mathbb{N}$, $s \in [0, 1]$.
 - $\|f \circ \Phi\|_{\mathcal{C}^s(B^n(\eta_1))} \lesssim_{\{s-1, s-2\}} \|f\|_{\mathcal{C}_{X_{J_0}}^s(B_{X_{J_0}}(x_0, \chi))}$, $s \in (0, \infty)$.

REMARK 4.8. The lack of optimality of Theorem 4.7 can be seen by comparing Theorem 4.7 (h) and Theorem 3.11 (vi); in the later one can estimate $\|Y_j\|_{\mathcal{C}^{s+1}}$ in terms of an $\{s\}$ -admissible constant, while in the former, one can only estimate $\|Y_j\|_{\mathcal{C}^s}$ in terms of the similar $\{s, s-1\}$ -admissible constants. Because of this, Theorem 4.7 “loses one derivative” and is not powerful enough to conclude necessary and sufficient results like Theorems 3.3 and 3.6.

REMARK 4.9. By comparing (h) and (j), we see that the functions $\tilde{c}_{j,k}^l$ have the same regularity as $Y_1, \dots, Y_n$. If one only knew the regularity of $Y_1, \dots, Y_n$, one could only conclude the regularity of $\tilde{c}_{j,k}^l$ for one fewer derivative. Similarly, (i) gives one more derivative regularity on b_k^l than we get from merely considering the regularity of $Y_1, \dots, Y_q$. In the second paper of this series, we will leverage this extra regularity to prove Theorem 3.11.

REMARK 4.10. Because the methods in this paper are based on ODEs, it is possible to prove versions of Theorem 4.7 for some function spaces other than $C^{m,s}$ or $\mathcal{C}^s$, with the same methods as in this paper. However, once we turn to the second paper in the series, where PDEs are used, we are forced to work with more specialized spaces—and that is the main motivation for using Zygmund spaces in this context.

REMARK 4.11. In the context of Lie groups, the coordinates given by Φ are sometimes called canonical coordinates of the first kind.

4.1 More on the assumptions. We further consider the constants $\eta > 0$ and $\delta_0 > 0$ which were introduced in Section 3.2. First we present two examples which show why these constants cannot be dispensed with in our results, and then we state a result which shows such constants always exist.

EXAMPLE 4.12. This example demonstrates the importance of η . Let $\mathfrak{M} = \mathbb{R}$, $q = 1$, $x_0 > 0$, and let $X_1 = x^2 \frac{\partial}{\partial x}$. In this case, η can be taken no larger than $1/x_0$ —i.e., X_1 satisfies $\mathcal{C}(x_0, x_0^{-1}, \mathbb{R})$ but does not satisfy $\mathcal{C}(x_0, \eta', \mathbb{R})$ for any $\eta' > x_0^{-1}$ (because the ODE $\dot{\gamma}(t) = \gamma(t)^2$, $\gamma(0) = x_0$ exists only for $t < \frac{1}{x_0}$). If Theorem 4.7 held with constants independent of η (and therefore independent of x_0), then we could conclude that X_1 satisfied $\mathcal{C}(x_0, \eta', \mathbb{R})$ for some η' independent of x_0 . This is because the condition $\mathcal{C}$ is invariant under a change of coordinates, and we can therefore check it in the coordinate system given by Φ in Theorem 4.7. This is a contradiction, showing η must play a role in Theorem 4.7.⁷

EXAMPLE 4.13. This example demonstrates the importance of δ_0 —and also shows its topological nature. The point of δ_0 is to ensure the map Φ in Theorem 4.7 is injective.⁸ Let $\mathfrak{M} = S^1$, $q = 1$, $x_0 \in S^1$, and let $X_1 = K \frac{\partial}{\partial \theta}$ for some large constant

⁷ For a similar example, one could take $\mathfrak{M} = (-\epsilon, \epsilon)$, $q = 1$, $x_0 = 0$, and $X_1 = \frac{\partial}{\partial x}$. Then, X satisfies $\mathcal{C}(0, \epsilon, (-\epsilon, \epsilon))$, but does not satisfy $\mathcal{C}(0, \eta', (-\epsilon, \epsilon))$ for any $\eta' > \epsilon$.

⁸ In fact, by inspecting the proof of Theorem 4.7, it is easy to see that one can prove similar results, independent of δ_0 , so long as one allows Φ to not be injective.

K . For this example, we must take $\delta_0 \leq 2\pi/K$. If the constants in Theorem 4.7 did not depend on δ_0 , they would also not depend on K . We could then conclude that δ_0 could be taken independent of K —this is because δ_0 is invariant under a change of coordinates and we can check it in the coordinate system given by Φ in Theorem 4.7—see also Proposition 4.14. This shows that δ_0 must play a role in Theorem 4.7.

Now we state a result which shows that such a δ_0 and η always exist for C^1 vector fields. Let $X_1, \dots, X_q$ be C^1 vector fields on a C^2 manifold $\mathfrak{M}$, and let X denote the list $X_1, \dots, X_q$.

PROPOSITION 4.14. • $\forall x_0 \in \mathfrak{M}, \exists \eta > 0$, such that X satisfies $\mathcal{C}(x_0, \eta, \mathfrak{M})$.

- Let $K \subseteq \mathfrak{M}$ be a compact set. Then, $\exists \delta_0 > 0$ such that $\forall \theta \in S^{q-1}$ if $x \in K$ is such that $\theta_1 X_1(x) + \dots + \theta_q X_q(x) \neq 0$, then $\forall r \in (0, \delta_0]$,

$$e^{r\theta_1 X_1 + \dots + r\theta_q X_q} x \neq x.$$

For the proof, see Section 9.5. Proposition 4.14 shows that there always exist η and δ_0 as in Section 3.2. However, the η and δ_0 guaranteed by Proposition 4.14 depend on the C^1 norms of $X_1, \dots, X_q$ in some fixed coordinate system, and this is not invariant under diffeomorphisms. It is important for some applications that η and δ_0 can be taken to be large in some settings even when the C^1 norms of $X_1, \dots, X_q$ are large. The next example gives a simple setting where this is the case.

EXAMPLE 4.15. Take $q = 1$, $\mathfrak{M} = \mathbb{R}$, $X_1 = K \frac{\partial}{\partial x}$, for any $K \in \mathbb{R} \setminus \{0\}$ (we think of K as large). Then one can take $\eta = \delta_0 = \infty$ in the assumptions in Section 3.2.

5 Wedge Products

Let Z be a one dimensional real vector space. For $x, y \in Z$, $x \neq 0$ we define $\frac{y}{x} \in \mathbb{R}$ by $\frac{y}{x} := \frac{\lambda(y)}{\lambda(x)}$ where $\lambda : Z \rightarrow \mathbb{R}$ is any nonzero linear functional. It is easy to see that $\frac{y}{x}$ is independent of the choice of λ .

This allows us to formulate a “coordinate free” version of Cramer’s rule. Let V be an n -dimensional vector space, so that $\bigwedge^n V$ is a one dimensional vector space. Let $x_1, \dots, x_n \in V$ be a basis for V . For any $y \in V$, we have

$$y = \frac{y \wedge x_2 \wedge x_3 \wedge \dots \wedge x_n}{x_1 \wedge x_2 \wedge \dots \wedge x_n} x_1 + \frac{x_1 \wedge y \wedge x_3 \wedge \dots \wedge x_n}{x_1 \wedge x_2 \wedge \dots \wedge x_n} x_2 + \dots + \frac{x_1 \wedge x_2 \wedge \dots \wedge x_{n-1} \wedge y}{x_1 \wedge x_2 \wedge \dots \wedge x_n} x_n. \quad (5.1)$$

Let M be a C^2 manifold of dimension n . Let $Y_1, \dots, Y_n$ be C^1 vector fields in on M . For another C^1 vector field Z , the Lie derivative of $Y_1 \wedge Y_2 \wedge \dots \wedge Y_n$ with respect Z is given by

$$\begin{aligned}\mathcal{L}_Z(Y_1 \wedge Y_2 \wedge \cdots \wedge Y_n) &= [Z, Y_1] \wedge Y_2 \wedge Y_3 \wedge \cdots \wedge Y_n + Y_1 \wedge [Z, Y_2] \wedge Y_3 \wedge \cdots \wedge Y_n \\ &\quad + \cdots + Y_1 \wedge Y_2 \wedge \cdots \wedge Y_{n-1} \wedge [Z, Y_n].\end{aligned}$$

Let $X_1, \dots, X_n$ be C^1 vector fields on M which span the tangent space near a point x_0 . Thus, near x_0 , we may define a real valued function by

$$\frac{Y_1 \wedge Y_2 \wedge \cdots \wedge Y_n}{X_1 \wedge X_2 \wedge \cdots \wedge X_n}.$$

The derivative of this function with respect to Z is exactly what one would expect as the next lemma shows.

LEMMA 5.1.

$$\begin{aligned}Z \frac{Y_1 \wedge Y_2 \wedge \cdots \wedge Y_n}{X_1 \wedge X_2 \wedge \cdots \wedge X_n} &= \frac{\mathcal{L}_Z(Y_1 \wedge Y_2 \wedge \cdots \wedge Y_n)}{X_1 \wedge X_2 \wedge \cdots \wedge X_n} \\ &\quad - \frac{Y_1 \wedge Y_2 \wedge \cdots \wedge Y_n}{X_1 \wedge X_2 \wedge \cdots \wedge X_n} \frac{\mathcal{L}_Z(X_1 \wedge X_2 \wedge \cdots \wedge X_n)}{X_1 \wedge X_2 \wedge \cdots \wedge X_n}.\end{aligned}$$

Proof. Let $\mathfrak{X} = X_1 \wedge X_2 \wedge \cdots \wedge X_n$ and $\mathfrak{Y} = Y_1 \wedge Y_2 \wedge \cdots \wedge Y_n$. Let ν be any C^1 n -form which is nonzero near x_0 , so that by definition

$$\frac{\mathfrak{Y}}{\mathfrak{X}} = \frac{\nu(\mathfrak{Y})}{\nu(\mathfrak{X})}.$$

Because ν is nonzero near x_0 (and the space of n -forms is one dimensional at each point), we may write $\mathcal{L}_Z \nu = f\nu$ for some continuous function f (near x_0); where here and in the rest of the paper $\mathcal{L}_Z$ denotes the Lie derivative with respect to Z . Using [Lee03, Proposition 18.9], we have

$$Z\nu(\mathfrak{Y}) = (\mathcal{L}_Z \nu)(\mathfrak{Y}) + \nu(\mathcal{L}_Z \mathfrak{Y}) = f\nu(\mathfrak{Y}) + \nu(\mathcal{L}_Z \mathfrak{Y}).$$

and similarly with $\mathfrak{Y}$ replaced by $\mathfrak{X}$. We conclude

$$\begin{aligned}Z \frac{\mathfrak{Y}}{\mathfrak{X}} &= Z \frac{\nu(\mathfrak{Y})}{\nu(\mathfrak{X})} = \frac{Z\nu(\mathfrak{Y})}{\nu(\mathfrak{X})} - \frac{\nu(\mathfrak{Y})}{\nu(\mathfrak{X})} \frac{Z\nu(\mathfrak{X})}{\nu(\mathfrak{X})} \\ &= \frac{f\nu(\mathfrak{Y}) + \nu(\mathcal{L}_Z \mathfrak{Y})}{\nu(\mathfrak{X})} - \frac{\nu(\mathfrak{Y})}{\nu(\mathfrak{X})} \frac{f\nu(\mathfrak{X}) + \nu(\mathcal{L}_Z \mathfrak{X})}{\nu(\mathfrak{X})} \\ &= \frac{\nu(\mathcal{L}_Z \mathfrak{Y})}{\nu(\mathfrak{X})} - \frac{\nu(\mathfrak{Y})}{\nu(\mathfrak{X})} \frac{\nu(\mathcal{L}_Z \mathfrak{X})}{\nu(\mathfrak{X})} = \frac{\mathcal{L}_Z \mathfrak{Y}}{\mathfrak{X}} - \frac{\mathfrak{Y}}{\mathfrak{X}} \frac{\mathcal{L}_Z \mathfrak{X}}{\mathfrak{X}},\end{aligned}$$

completing the proof. $\square$

6 Densities

Let $\chi \in (0, \xi]$ be as in Theorem 4.7. In many applications, one is given a density on $B_{X_{j_0}}(x_0, \chi)$ and it is of interest to measure certain sets with respect to this density. For a quick introduction to the basics of densities, we refer the reader to Guillemin's lecture notes [Gui08].

Let ν be a C^1 density on $B_{X_{J_0}}(x_0, \chi)$. Suppose

$$\mathcal{L}_{X_j}\nu = f_j\nu, \quad 1 \leq j \leq n, \quad f_j \in C(B_{X_{J_0}}(x_0, \chi)). \quad (6.1)$$

Our goal is to understand $\Phi^*\nu$ and $\nu(B_X(x_0, \xi_2))$ where Φ and ξ_2 are as in Theorem 4.7.

DEFINITION 6.1. We say C is a $0; \nu$ -admissible constant if C is a 0-admissible constant which is also allowed to depend on upper bounds for $\|f_j\|_{C(B_{X_{J_0}}(x_0, \chi))}$, $1 \leq j \leq n$.

DEFINITION 6.2. We say C is a $1; \nu$ -admissible constant if C is a 1-admissible constant which is also allowed to depend on upper bounds for $\|f_j\|_{C(B_{X_{J_0}}(x_0, \chi))}$, $1 \leq j \leq n$.

DEFINITION 6.3. For $m_1, m_2 \in \mathbb{Z}$, $s \in [0, 1]$ if we say C is an $\langle m_1, m_2, s; \nu \rangle$ -admissible constant, it means that we assume $f_j \in C_{X_{J_0}}^{m_1, s}(B_{X_{J_0}}(x_0, \chi))$, and C is an $\langle m_1, m_2, s \rangle$ -admissible constant which is also allowed to depend on upper bounds for $\|f_j\|_{C_{X_{J_0}}^{m_1, s}(B_{X_{J_0}}(x_0, \chi))}$, $1 \leq j \leq n$.

DEFINITION 6.4. For $s_1 > 0$, $s_2 \in \mathbb{R}$, if we say C is an $\{s_1, s_2; \nu\}$ -admissible constant, it means that we assume $f_j \in \mathcal{C}_{X_{J_0}}^{s_1}(B_{X_{J_0}}(x_0, \chi))$, and C is an $\{s_1, s_2\}$ -admissible constant which is also allowed to depend on upper bounds for $\|f_j\|_{\mathcal{C}_{X_{J_0}}^{s_1}(B_{X_{J_0}}(x_0, \chi))}$, $1 \leq j \leq n$. For $s_1 \leq 0$, $s_2 \in \mathbb{R}$, if we say C is an $\{s_1, s_2; \nu\}$ -admissible constant, it means C is an $\{s_1, s_2\}$ -admissible constant which is also allowed to depend on upper bounds for $\|f_j\|_{C(B_{X_{J_0}}(x_0, \chi))}$, $1 \leq j \leq n$.

We write $A \lesssim_{0; \nu} B$ for $A \leq CB$ where C is a positive $0; \nu$ -admissible constant, and write $A \approx_{0; \nu} B$ for $A \lesssim_{0; \nu} B$ and $B \lesssim_{0; \nu} A$. We define $\lesssim_{1; \nu}$, $\approx_{1; \nu}$, $\lesssim_{\langle m_1, m_2, s; \nu \rangle}$, $\approx_{\langle m_1, m_2, s; \nu \rangle}$, $\lesssim_{\{s_1, s_2; \nu\}}$, and $\approx_{\{s_1, s_2; \nu\}}$ similarly.

To help understand ν , we use a distinguished density ν_0 on $B_{X_{J_0}}(x_0, \chi)$:

$$\nu_0(Z_1, \dots, Z_n) := \left| \frac{Z_1 \wedge Z_2 \wedge \dots \wedge Z_n}{X_1 \wedge X_2 \wedge \dots \wedge X_n} \right|, \quad (6.2)$$

note that ν_0 is defined since $X_1 \wedge X_2 \wedge \dots \wedge X_n$ is never zero on $B_{X_{J_0}}(x_0, \chi)$ by Theorem 4.7 (a); ν_0 is clearly a density.

Theorem 6.5. There exists $g \in C(B_{X_{J_0}}(x_0, \chi))$ such that $\nu = g\nu_0$ and

- (i) $g(x) \approx_{0; \nu} g(x_0) = \nu(X_1, \dots, X_n)(x_0)$, $\forall x \in B_{X_{J_0}}(x_0, \chi)$. In particular, g always has the same sign, and is either never zero or always zero.
- (ii) We have the following regularity on g :
 - For $m \in \mathbb{N}$, $s \in [0, 1]$, we have $\|g\|_{C_{X_{J_0}}^{m, s}(B_{X_{J_0}}(x_0, \chi))} \lesssim_{\langle m-1, m-1, s; \nu \rangle} |\nu(X_1, \dots, X_n)(x_0)|$.
 - For $s > 0$, we have $\|g\|_{\mathcal{C}_{X_{J_0}}^s(B_{X_{J_0}}(x_0, \chi))} \lesssim_{\{s-1, s-1; \nu\}} |\nu(X_1, \dots, X_n)(x_0)|$.

Define $h \in C^1(B^n(\eta_1))$ by $\Phi^*\nu = h\sigma_{\text{Leb}}$, where σ_{Leb} denotes the usual Lebesgue density on $\mathbb{R}^n$.

- (iii) $h(t) \approx_{0;\nu} \nu(X_1, \dots, X_n)(x_0)$, $\forall t \in B^n(\eta_1)$. In particular, h always has the same sign and is either never zero or always zero.
- (iv) We have the following regularity on h :
 - For $m \in \mathbb{N}$, $s \in [0, 1]$, $\|h\|_{C^{m,s}(B^n(\eta_1))} \lesssim_{\langle m, m-1, s; \nu \rangle} |\nu(X_1, \dots, X_n)(x_0)|$.
 - For $s > 0$, $\|h\|_{\mathcal{C}^s(B^n(\eta_1))} \lesssim_{\{s, s-1; \nu\}} |\nu(X_1, \dots, X_n)(x_0)|$.

COROLLARY 6.6. Let ξ_2 be as in Theorem 4.7. Then,

$$\nu(B_{X_{j_0}}(x_0, \xi_2)) \approx_{1;\nu} \nu(B_X(x_0, \xi_2)) \approx_{1;\nu} \nu(X_1, \dots, X_n)(x_0), \quad (6.3)$$

and therefore,

$$\begin{aligned} |\nu(B_{X_{j_0}}(x_0, \xi_2))| &\approx_{1;\nu} |\nu(B_X(x_0, \xi_2))| \approx_{1;\nu} |\nu(X_1, \dots, X_n)(x_0)| \approx_0 \\ &\max_{(j_1, \dots, j_n) \in \mathcal{I}(n, q)} |\nu(X_{j_1}, \dots, X_{j_n})(x_0)|. \end{aligned} \quad (6.4)$$

7 Scaling and Other Consequences

The main results of this series have two facets:

- (Smoothness) They provide a coordinate system in which given C^1 vector fields have an optimal degree of smoothness.
- (Scaling) They provide a coordinate system in which given vector fields are normalized in a way which is useful for applying techniques from analysis.

In both cases, the results are in many ways optimal: they provide necessary and sufficient, diffeomorphic invariant conditions under which one can obtain such coordinate charts. In this section, we describe these two facets.

When viewed as providing a coordinate system in which vector fields have an optimal level of smoothness, these results seem to be of a new type. When viewed as scaling maps, these results take their roots in the quantitative study of sub-Riemannian (aka Carnot–Carathéodory) geometry initiated by Nagel, Stein, and Wainger [NSW85]. Since Nagel, Stein, and Wainger’s original work, these ideas have had a significant impact on various questions in harmonic analysis (see the discussion at the end of Chapter 2 of [Str14] for a detailed history of these ideas). Following Nagel, Stein, and Wainger’s work, Tao and Wright [TW03] generalized Nagel, Stein, and Wainger’s ideas and provided a new approach to proving their results. In [Str11], the second author combined these two approaches to prove results in more general settings; these more general results have already had several applications, for example [SS11, Str12, SS13, SS12, Str17, Str14, Gre15, Sto14].

7.1 Classical sub-Riemannian geometries and the work of Nagel, Stein, and Wainger.

In this section, we describe the foundational work of Nagel, Stein, and Wainger [NSW85], and see how it is a special case of Theorem 4.7. This provides the simplest non-trivial setting where the results in this paper can be seen as providing scaling maps adapted to a sub-Riemannian geometry. In Section 7.3, we generalize these results to more general geometries.

Let $X_1, \dots, X_q$ be C^∞ vector fields on an open set $\Omega \subseteq \mathbb{R}^n$; we assume $X_1, \dots, X_q$ span the tangent space at every point of Ω . To each X_j assign a formal degree $d_j \in [1, \infty)$. We assume

$$[X_j, X_k] = \sum_{d_l \leq d_j + d_k} c_{j,k}^l X_l, \quad c_{j,k}^l \in C^\infty(\Omega). \quad (7.1)$$

We write (X, d) for the list $(X_1, d_1), \dots, (X_q, d_q)$ and for $\delta > 0$ write $\delta^d X$ for the list of vector fields $\delta^{d_1} X_1, \dots, \delta^{d_q} X_q$. The sub-Riemannian ball associated to (X, d) centered at $x_0 \in \Omega$ of radius $\delta > 0$ is defined by

$$B_{(X,d)}(x_0, \delta) := B_{\delta^d X}(x_0, 1),$$

where the later ball is defined by (2.2). $B_{(X,d)}(x_0, \delta)$ is an open subset of Ω . It is easy to see that the balls $B_{(X,d)}(x, \delta)$ are metric balls.

Define, for $x \in \Omega$, $\delta \in (0, 1]$,

$$\Lambda(x, \delta) := \max_{j_1, \dots, j_n \in \{1, \dots, q\}} \left| \det \left(\delta^{d_{j_1}} X_{j_1}(x) \cdots \delta^{d_{j_n}} X_{j_n}(x) \right) \right|.$$

For each $x \in \Omega$, $\delta \in (0, 1]$, pick $j_1 = j_1(x, \delta), \dots, j_n = j_n(x, \delta)$ so that

$$\left| \det \left(\delta^{d_{j_1}} X_{j_1}(x) \cdots \delta^{d_{j_n}} X_{j_n}(x) \right) \right| = \Lambda(x, \delta).$$

For this choice of $j_1 = j_1(x, \delta), \dots, j_n = j_n(x, \delta)$, set

$$\Phi_{x,\delta}(t_1, \dots, t_n) := \exp \left(t_1 \delta^{d_{j_1}} X_{j_1} + \cdots + t_n \delta^{d_{j_n}} X_{j_n} \right) x.$$

Theorem 7.1 ([NSW85]). *Fix a compact set $\mathcal{K} \Subset \Omega$.⁹ In what follows, we write $A \lesssim B$ for $A \leq CB$ where C is a positive constant which may depend on $\mathcal{K}$, but does not depend on the particular point $x \in \mathcal{K}$ or the scale $\delta \in (0, 1]$. There exist $\eta_1, \xi_0 \approx 1$, such that $\forall x \in \mathcal{K}$,*

- (i) $\sigma_{\text{Leb}}(B_{(X,d)}(x, \delta)) \approx \Lambda(x, \delta)$, $\forall \delta \in (0, \xi_0]$.
- (ii) $\sigma_{\text{Leb}}(B_{(X,d)}(x, 2\delta)) \lesssim \sigma_{\text{Leb}}(B_{(X,d)}(x, \delta))$, $\forall \delta \in (0, \xi_0/2]$.
- (iii) $\forall \delta \in (0, 1]$, $\Phi_{x,\delta}(B^n(\eta_1)) \subseteq \Omega$ is open and $\Phi_{x,\delta} : B^n(\eta_1) \rightarrow \Phi_{x,\delta}(B^n(\eta_1))$ is a C^∞ diffeomorphism.
- (iv) $|\det d\Phi_{x,\delta}(t)| \approx \Lambda(x, \delta)$, $\forall t \in B^n(\eta_1)$.

⁹ Here, and in the rest of the paper, we write $\mathcal{K} \Subset \Omega$ to mean that $\mathcal{K}$ is a relatively compact subset of Ω .

- (v) $B_{(X,d)}(x, \xi_0 \delta) \subseteq \Phi_{x,\delta}(B^n(\eta_1)) \subseteq B_{(X,d)}(x, \delta)$, $\forall \delta \in (0, 1]$.
 (vi) Let $Y_j^{x,\delta} := \Phi_{x,\delta}^* \delta^{d_j} X_j$, so that $Y_j^{x,\delta}$ is a C^∞ vector field on $B^n(\eta_1)$. We have

$$\left\| Y_j^{x,\delta} \right\|_{C^m(B^n(\eta_1); \mathbb{R}^n)} \lesssim 1, \quad \forall m \in \mathbb{N},$$

where the implicit constant depends on m , by not on $x \in \mathcal{K}$ or $\delta \in (0, 1]$. Finally, $Y_1^{x,\delta}(u), \dots, Y_q^{x,\delta}(u)$ span $T_u B^n(\eta_1)$, uniformly in x, δ , and u , in the sense that

$$\max_{j_1, \dots, j_n \in \{1, \dots, q\}} \inf_{u \in B^n(\eta_1)} \left| \det \left(Y_{j_1}^{x,\delta}(u) | \dots | Y_{j_n}^{x,\delta}(u) \right) \right| \approx 1.$$

Proof. This result is a special case of Theorem 7.6, below. To see this, for $\delta \in (0, 1]$ we multiply both sides of (7.1) by $\delta^{d_j+d_k}$ to obtain

$$[\delta^{d_j} X_j, \delta^{d_k} X_k] = \sum_{d_l \leq d_j + d_k} \delta^{d_j + d_k - d_l} c_{j,k}^l \delta^{d_l} X_l,$$

so that if we set

$$X_j^\delta := \delta^{d_j} X_j, \quad c_{j,k}^{l,\delta} := \begin{cases} \delta^{d_j + d_k - d_l} c_{j,k}^l, & d_l \leq d_j + d_k, \\ 0, & \text{otherwise,} \end{cases}$$

then we have

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

Furthermore, $c_{j,k}^{l,\delta} \in C^\infty$ and $X_l^\delta \in C^\infty$ uniformly in δ . From here it is straightforward to verify that $X_1^\delta, \dots, X_q^\delta$ satisfy all the hypotheses of Theorem 7.6; in the application of Theorem 7.6, we replace Ω with Ω' where $\mathcal{K} \subseteq \Omega' \subseteq \Omega$. $\square$

REMARK 7.2. It is easy to see that the balls $B_{(X,d)}(x, \delta)$ are metric balls.¹⁰ Theorem 7.1 (ii) is the main estimate needed to show these balls (when paired with σ_{Leb}) form a space of homogeneous type. Thus, one can obtain a theory of singular integrals associated with these balls. Such singular integrals have a long history and have proven to be quite useful in a variety of contexts. The history of these ideas is detailed at the end of [Str14, Chapter 2].

7.1.1 Hörmander's condition. The main way that Theorem 7.1 arises is via vector fields which satisfy Hörmander's condition. Suppose $V_1, \dots, V_r$ are C^∞ vector fields on an open set $\Omega \subseteq \mathbb{R}^n$. We assume that $V_1, \dots, V_r$ satisfy Hörmander's condition of order m on Ω . I.e., we assume that the finite list of vector fields

$$V_1, \dots, V_r, \dots, [V_i, V_j], \dots, [V_i, [V_j, V_k]], \dots, \dots, \text{commutators of order } m,$$

span the tangent space at every point of Ω .

¹⁰ This uses that $d_j \geq 1$, $\forall j$. If $d_j \in (0, \infty)$, they are quasi-metric balls.

To each $V_1, \dots, V_r$, we assign the formal degree 1. If Z has formal degree e , we assign to $[V_j, Z]$ the formal degree $e + 1$. Let $(X_1, d_1), \dots, (X_q, d_q)$ denote the finite list of vector fields with formal degree $d_j \leq m$. Hörmander's condition implies $X_1, \dots, X_q$ span the tangent space at every point of Ω .

We claim that (7.1) holds, and therefore Theorem 7.1 applies to $(X_1, d_1), \dots, (X_q, d_q)$. Indeed, if $d_j + d_k \leq m$ we have

$$[X_j, X_k] = \sum_{d_l = d_j + d_k} c_{j,k}^l X_l,$$

where $c_{j,k}^l$ are constants by the Jacobi identity. If $d_j + d_k > m$ then, since $X_1, \dots, X_q$ span the tangent space at every point, we have

$$[X_j, X_k] = \sum_{l=1}^q c_{j,k}^l X_l = \sum_{d_l \leq d_j + d_k} c_{j,k}^l X_l, \quad c_{j,k}^l \in C^\infty(\Omega).$$

Thus, (7.1) holds and Theorem 7.1 applies.

Let $\mathcal{K} \Subset \Omega$ be a compact set. Applying Theorem 7.1, for $\delta \in (0, 1]$, $x \in \mathcal{K}$, we obtain $\eta_1 > 0$ and $\Phi_{x,\delta} : B^n(\eta_1) \rightarrow B_{(X,d)}(x, \delta)$ as in that theorem. Set $V_j^{x,\delta} := \Phi_{x,\delta}^* \delta V_j$, $1 \leq j \leq r$.

If $d_k = l$, then

$$X_k = [V_{j_1}, [V_{j_2}, \dots, [V_{j_{l-1}}, V_{j_l}] \dots]],$$

and so

$$\Phi_{x,\delta}^* \delta^{d_k} X_k = \Phi_{x,\delta}^* [\delta V_{j_1}, [\delta V_{j_2}, \dots, [\delta V_{j_{l-1}}, \delta V_{j_l}] \dots]] = [V_{j_1}^{x,\delta}, [V_{j_2}^{x,\delta}, \dots, [V_{j_{l-1}}^{x,\delta}, V_{j_l}^{x,\delta}] \dots]].$$

Theorem 7.1 implies that the vector fields $\Phi_{x,\delta}^* \delta^{d_k} X_k$ are smooth and span the tangent space, uniformly for $x \in \mathcal{K}$, $\delta \in (0, 1]$. We conclude that the vector fields $V_1^{x,\delta}, \dots, V_r^{x,\delta}$ are smooth and satisfy Hörmander's condition, uniformly for $x \in \mathcal{K}$, $\delta \in (0, 1]$. In short, the map $\Phi_{x,\delta}^*$ takes $\delta V_1, \dots, \delta V_r$ to $V_1^{x,\delta}, \dots, V_r^{x,\delta}$ which satisfy Hörmander's condition “uniformly”; i.e., it takes the case of δ small and “rescales” it to the case $\delta = 1$.

REMARK 7.3. In the above, we multiplied $V_1, \dots, V_r$ all by the same small number δ . Similar results hold (with the same proofs) for $\delta_1 V_1, \dots, \delta_r V_r$ where $\delta_1, \dots, \delta_r$ are small, provided they are “weakly-comparable.” I.e., provided $\exists N, \kappa$ such that $\delta_j^N \leq \kappa \delta_k$, for all j, k . This was first noted and used by Tao and Wright [TW03]. See [Str11, Section 5.2.1] for further details.

REMARK 7.4. It is possible for (7.1) to hold (for a sufficiently large m) even if $V_1, \dots, V_r$ do not satisfy Hörmander's condition. In this case, with the same proof one can obtain similar results; however, now the ball $B_{(X,d)}(x, \delta)$ lies on an injectively immersed submanifold of $\mathbb{R}^n$ as discussed in Proposition 3.1. An important setting where this arises is when $V_1, \dots, V_r$ are real analytic; see [Str14, Section 2.15.5] for details.

7.2 Multi-parameter balls. In a generalization of the work of Nagel, Stein, and Wainger, the second author studied multi-parameter sub-Riemannian balls in [Str11]. The main result of [Str11] is a special case of Theorems 4.7 and 6.5 and Corollary 6.6. We refer the reader to [Str11] for the detailed assumptions used in that paper, which are very similar to the assumptions of Theorem 4.7. We give a few comments here to help the reader understand how the main result of [Str11] (namely [Str11, Theorem 4.1]) is a special case of the results in this paper.

The main differences between [Str11, Theorem 4.1] and the setting of this paper are:

- $\mathfrak{M}$ is taken to be an open subset of $\mathbb{R}^N$ in [Str11].
- In [Str11], the various kinds of admissible constants are allowed to depend on upper bounds for quantities like $\|X_j\|_{C^m}$. This quantity is not invariant under diffeomorphisms, and the norm is defined in terms of the fixed standard coordinate system on $\mathbb{R}^N$.
- Instead of an abstract density as is used in Theorem 6.5 and Corollary 6.6, [Str11] uses the usual Lebesgue measure on submanifolds of $\mathbb{R}^N$.
- In [Str11], the existence of δ_0 is not assumed. Instead, one uses bounds on $\|X_j\|_{C^1}$ to prove that such a δ_0 exists (as in Proposition 4.14). This process is not invariant under diffeomorphisms.
- The constants in Theorem 4.7 have better dependence on various quantities than they do in [Str11, Theorem 4.1]. For example, the methods in [Str11] do not imply that η_1 is a 1-admissible constant.
- In [Str11], only the spaces C_X^m (and not $C_X^{m,s}$ or $\mathcal{C}_X^s$) were used.

We include a lemma, whose straightforward proof we omit, which will allow the reader to more easily translate the results of [Str11] into the language of this paper. For an $N \times n$ matrix we write $\det_{n \times n} B$ to be the vector consisting of determinants of $n \times n$ submatrices of B .

LEMMA 7.5. *Let L be an n -dimensional injectively immersed submanifold of $\mathbb{R}^N$, and give L the induced Riemannian metric. Let ν denote the Riemannian volume density on L . For vector fields $Z_1, \dots, Z_n$ on $\mathbb{R}^N$ which are tangent to L , let Z denote the $N \times n$ matrix whose columns are $Z_1, \dots, Z_n$. Then,*

$$\left| \det_{n \times n} Z \right| = \nu(Z_1, \dots, Z_n).$$

Furthermore, if $\Phi : B^n(\eta) \rightarrow L \subseteq \mathbb{R}^N$, and if $\Phi^* \nu = h(t) \sigma_{\text{Leb}}$, then we have

$$h(t) = \left| \det_{n \times n} d\Phi(t) \right|,$$

where $d\Phi(t)$ is computed by thinking of Φ as a map $B^n(\eta) \rightarrow \mathbb{R}^N$.

Using this lemma and the above remarks, [Str11, Theorem 4.1] follows easily from the results in this paper. We refer the reader to [Str11, Str14, SS11, Str12, SS13, SS12, Str17] for examples of how these ideas can be used as scaling maps.

7.3 Generalized sub-Riemannian geometries. The results described in Section 7.1 concern the classical setting of sub-Riemannian geometry. When applied to partial differential equations defined by vector fields, this is the geometry which arises in the important case of *maximally hypoelliptic* operators. Maximal hypoellipticity is a far reaching generalization of ellipticity, which was first introduced (implicitly) by Folland and Stein [FS74]; see [Str14, Chapter 2] for a discussion of these ideas as well as a detailed history. When one moves beyond the setting of maximal hypoellipticity, other more general sub-Riemannian geometries can arise. These are defined by choosing different vector fields at each scale. A particularly transparent setting where this arises is in the work of Charpentier and Dupain on the Bergman and Szegő projections [CD14]. The theory in this paper allows us to easily understand what properties one requires on these vector fields so that the induced quasi-metrics give rise to a space of homogeneous type; furthermore, our theory provides generalized scaling maps adapted to these geometries. See Section 7.5 for some further comments on the relationship between the results in this paper and several complex variables.

Fix an open set $\Omega \subseteq \mathbb{R}^n$, and for each $\delta \in (0, 1]$, let $X^\delta = X_1^\delta, \dots, X_q^\delta$ be a list of C^1 vector fields on Ω , which span the tangent space at every point. For $x \in \Omega$, $\delta \in (0, 1]$ set $B(x, \delta) := B_{X^\delta}(x, 1)$, where $B_{X^\delta}(x, 1)$ is defined by (2.2). Our goal is to give conditions on X^δ so that the balls $B(x, \delta)$, when paired with Lebesgue measure on Ω (denoted σ_{Leb}), locally form a space of homogeneous type (see [Ste93] for the definition we are using of a space of homogeneous type). The conditions we give can be thought of as infinitesimal versions of the axioms of a space of homogeneous type. In what follows, we write X^δ for the column vector of vector fields $[X_1^\delta, \dots, X_q^\delta]^\top$. Because of this, if we are given a matrix $A : \Omega \rightarrow \mathbb{M}^{q \times q}$, it makes sense to consider $A(x)X^\delta(x)$ which again gives a column vector of vector fields on Ω .

We assume:

- (I) $\forall \delta \in (0, 1]$, $x \in \Omega$, we have $\text{span}\{X_1^\delta(x), \dots, X_q^\delta(x)\} = T_x\Omega$.
- (II) $\sup_{\delta \in (0, 1]} \|X_j^\delta\|_{C^1(\Omega; \mathbb{R}^n)} < \infty$.
- (III) $X_j^\delta \rightarrow 0$, as $\delta \rightarrow 0$, uniformly on compact subsets of Ω .
- (IV) $\forall 0 < \delta_1 \leq \delta_2 \leq 1$, $X^{\delta_1} = T_{\delta_1, \delta_2} X^{\delta_2}$, where $T_{\delta_1, \delta_2} \in L^\infty(\Omega; \mathbb{M}^{q \times q})$, and $\|T_{\delta_1, \delta_2}\|_{L^\infty(\Omega; \mathbb{M}^{q \times q})} \leq 1$.
- (V) $\exists B_1, B_2 \in (1, \infty)$, $b_1, b_2 \in (0, 1)$, such that $\forall \delta \in (0, 1/B_1]$, $\exists S_\delta \in L^\infty(\Omega; \mathbb{M}^{q \times q})$ and $\forall \delta \in (0, 1/B_2]$, $\exists R_\delta \in L^\infty(\Omega; \mathbb{M}^{q \times q})$ with $S_\delta X^{B_1\delta} = X^\delta$, $R_\delta X^\delta = X^{B_2\delta}$, and

$$\sup_{0 < \delta \leq 1/B_1} \|S_\delta\|_{L^\infty(\Omega; \mathbb{M}^{q \times q})} \leq b_1, \quad \sup_{0 < \delta \leq 1/B_2} \|R_\delta\|_{L^\infty(\Omega; \mathbb{M}^{q \times q})} \leq b_2^{-1}.$$

- (VI) $\forall \delta \in (0, 1]$, $[X_j^\delta, X_k^\delta] = \sum_{l=1}^q c_{j,k}^{l,\delta} X_l^\delta$, where $c_{j,k}^{l,\delta} \in C(\Omega)$ and $\forall m \in \mathbb{N}$

$$\sup_{\delta \in (0, 1], x \in \Omega} \|c_{j,k}^{l,\delta}\|_{C_{X^\delta}^m(B(x, \delta))} < \infty.$$

Define, for $x \in \Omega$, $\delta \in (0, 1]$,

$$\Lambda(x, \delta) := \max_{j_1, \dots, j_n \in \{1, \dots, q\}} \left| \det \left(X_{j_1}^\delta(x) | \cdots | X_{j_n}^\delta(x) \right) \right|.$$

For each $x \in \Omega$, $\delta \in (0, 1]$, pick $j_1 = j_1(x, \delta), \dots, j_n = j_n(x, \delta) \in \{1, \dots, q\}$ so that

$$\left| \det \left(X_{j_1}^\delta(x) | \cdots | X_{j_n}^\delta(x) \right) \right| = \Lambda(x, \delta),$$

and set (for this choice of $j_1 = j_1(x, \delta), \dots, j_n = j_n(x, \delta)$),

$$\Phi_{x, \delta}(t_1, \dots, t_n) = \exp \left(t_1 X_{j_1}^\delta + \cdots + t_n X_{j_n}^\delta \right) x.$$

Theorem 7.6.

- (i) $B(x, \delta_1) \subseteq B(x, \delta_2)$, $\forall x \in \Omega$, $0 < \delta_1 \leq \delta_2 \leq 1$.
- (ii) $\bigcap_{\delta \in (0, 1]} \overline{B(x, \delta)} = \{x\}$, $\forall x \in \Omega$.
- (iii) $B(x, \delta) \cap B(y, \delta) \neq \emptyset \Rightarrow B(y, \delta) \subseteq B(x, C\delta)$, $\forall \delta \in (0, 1/C]$, where $C = B_1^k$ and k is chosen so that $b_1^k \leq \frac{1}{3}$.
- (iv) For each $U \subseteq \Omega$ with U open, $\delta \in (0, 1]$, the map $x \mapsto \sigma_{\text{Leb}}(U \cap B(x, \delta))$ is continuous.

Fix a compact set $\mathcal{K} \subseteq \Omega$. In what follows we write $A \lesssim B$ for $A \leq CB$ where C is a positive constant which may depend on $\mathcal{K}$, but does not depend on the particular point $x \in \mathcal{K}$ or the scale $\delta \in (0, 1]$. We write $A \approx B$ for $A \lesssim B$ and $B \lesssim A$. There exist $\eta_1, \xi_0 \approx 1$ such that $\forall x \in \mathcal{K}$:

- (v) $\sigma_{\text{Leb}}(B(x, \delta)) \approx \Lambda(x, \delta)$, $\forall \delta \in (0, \xi_0]$.
- (vi) $\sigma_{\text{Leb}}(B(x, 2\delta)) \lesssim \sigma_{\text{Leb}}(B(x, \delta))$, $\forall \delta \in (0, \xi_0/2]$.
- (vii) $\forall \delta \in (0, 1]$, $\Phi_{x, \delta}(B^n(\eta_1)) \subseteq \Omega$ is open and $\Phi_{x, \delta} : B^n(\eta_1) \rightarrow \Phi_{x, \delta}(B^n(\eta_1))$ is a C^2 diffeomorphism.
- (viii) $|\det d\Phi_{x, \delta}(t)| \approx \Lambda(x, \delta)$, $\forall t \in B^n(\eta_1)$, $\delta \in (0, 1]$.
- (ix) $B(x, \xi_0\delta) \subseteq \Phi_{x, \delta}(B^n(\eta_1)) \subseteq B(x, \delta)$, $\forall \delta \in (0, 1]$.
- (x) Let $Y_j^{x, \delta} := \Phi_{x, \delta}^* X_j^\delta$, so that $Y_j^{x, \delta}$ is a vector field on $B^n(\eta_1)$. Then $Y_j^{x, \delta} \in C^\infty(B^n(\eta_1); \mathbb{R}^n)$ and

$$\|Y_j^{x, \delta}\|_{C^m(B^n(\eta_1); \mathbb{R}^n)} \lesssim 1, \quad \forall m \in \mathbb{N},$$

where the implicit constant may depend on m , but does not depend on $x \in \mathcal{K}$ or $\delta \in (0, 1]$. Furthermore, $Y_1^{x, \delta}(u), \dots, Y_q^{x, \delta}(u)$ span $T_u B^n(\eta_1)$, uniformly in x , δ , and u in the sense that

$$\max_{j_1, \dots, j_n \in \{1, \dots, q\}} \inf_{u \in B^n(\eta_1)} \left| \det \left(Y_{j_1}^{x, \delta}(u) | \cdots | Y_{j_n}^{x, \delta}(u) \right) \right| \approx 1.$$

Proof. To facilitate the proof, we introduce some new notation. For $y \in \Omega$, $y \in B(x, \delta) = B_{X^\delta}(x, 1)$ if and only if $\exists \gamma : [0, 1] \rightarrow \Omega$, $\gamma(0) = x$, $\gamma(1) = y$, $\gamma'(t) = \langle a(t), X^\delta(\gamma(t)) \rangle$, where $a \in L^\infty([0, 1]; \mathbb{R}^q)$ with $\|a\|_{L^\infty([0, 1]; \mathbb{R}^q)} < 1$, we have identified X^δ with the vector of vector fields $X^\delta = (X_1^\delta, \dots, X_q^\delta)$, and $\langle \cdot, \cdot \rangle$ denotes the usual inner product on $\mathbb{R}^q$.

(i): Let $0 < \delta_1 \leq \delta_2 \leq 1$. Take $y \in B(x, \delta_1)$ so that $\exists \gamma : [0, 1] \rightarrow \Omega$, $\gamma(0) = x$, $\gamma(1) = y$, $\gamma'(t) = \langle a(t), X^{\delta_1}(\gamma(t)) \rangle$, $\|a\|_{L^\infty([0, 1]; \mathbb{R}^q)} < 1$. We have

$$\begin{aligned} \gamma'(t) &= \langle a(t), X^{\delta_1}(\gamma(t)) \rangle = \langle a(t), T_{\delta_1, \delta_2}(\gamma(t)) X^{\delta_2}(\gamma(t)) \rangle \\ &= \langle T_{\delta_1, \delta_2}(\gamma(t))^\top a(t), X^{\delta_2}(\gamma(t)) \rangle. \end{aligned}$$

Since $\|T_{\delta_1, \delta_2}(\gamma(t))^\top a\|_{L^\infty([0, 1]; \mathbb{R}^q)} \leq \|a\|_{L^\infty([0, 1]; \mathbb{R}^q)} < 1$, this proves $y \in B(x, \delta_2)$, completing the proof of (i).

(ii) follows from the hypothesis (III).

(iii): Suppose $B(x, \delta) \cap B(y, \delta) \neq \emptyset$. This is equivalent to $B_{X^\delta}(x, 1) \cap B_{X^\delta}(y, 1) \neq \emptyset$. Since the balls $B_{X^\delta}(x, \cdot)$ are metric balls, this implies $B(y, \delta) = B_{X^\delta}(y, 1) \subseteq B_{X^\delta}(x, 3)$. Thus it suffices to show $B_{X^\delta}(x, 3) \subseteq B(x, C\delta)$. Suppose $z \in B_{X^\delta}(x, 3)$, so that $\exists \gamma : [0, 1] \rightarrow \Omega$, $\gamma(0) = x$, $\gamma(1) = z$, $\gamma'(t) = \langle a(t), 3X^\delta(\gamma(t)) \rangle$, where $\|a\|_{L^\infty([0, 1]; \mathbb{R}^q)} < 1$.

Take k so large that $b_1^k \leq \frac{1}{3}$. Then, for $\delta \in (0, B_1^{-k}]$,

$$\gamma'(t) = \langle a(t), A(t) X^{B_1^k \delta}(\gamma(t)) \rangle = \langle A(t)^\top a(t), X^{B_1^k \delta}(\gamma(t)) \rangle,$$

where

$$A(t) = 3S_\delta(\gamma(t))S_{B_1\delta}(\gamma(t)) \cdots S_{B_1^{k-1}\delta}(\gamma(t)).$$

Since $\|A\|_{L^\infty([0, 1]; \mathbb{M}^{q \times q})} \leq 3b_1^k \leq 1$, it follows that $\|A^\top a\|_{L^\infty([0, 1]; \mathbb{R}^q)} \leq \|a\|_{L^\infty([0, 1]; \mathbb{R}^q)} < 1$, and therefore $z = \gamma(1) \in B(x, B_1^k \delta) = B(x, C\delta)$, completing the proof of (iii).

(iv) follows from standard ODE results.

For the remaining parts, the goal is to apply Theorems 4.7 and 6.5 and Corollary 6.6 to the list of vector fields X^δ (with $\nu = \sigma_{\text{Leb}}$ and $\xi = 1$). Take $\eta \in (0, 1]$, depending on $\mathcal{K}$ and upper bounds for $\|X_j^\delta\|_{C^1(\Omega)}$, so that $\forall x \in \mathcal{K}$, $X_1^\delta, \dots, X_q^\delta$ satisfy $\mathcal{C}(x, \eta, \Omega)$. Note that η can be chosen independent of $x \in \mathcal{K}$ and $\delta \in (0, 1]$. Take $\delta_0 > 0$ as in Proposition 4.14 when applied to $X_1^\delta, \dots, X_q^\delta$, with $\mathfrak{M} = \Omega$. It can be seen from the proof of Proposition 4.14 that δ_0 can be chosen independent of $\delta \in (0, 1]$. Finally, note that $\mathcal{L}_{X_j^\delta} \nu = \text{div}(X_j^\delta) \nu =: f_j^\delta \nu$, where $\sup_{\delta \in (0, 1]} \|f_j^\delta\|_{C(\Omega)} < \infty$.

Using the above choices, all of the hypotheses of Theorems 4.7, 6.5 and Corollary 6.6 hold for $x_0 \in \mathcal{K}$ with $X_1, \dots, X_q$ replaced by $X_1^\delta, \dots, X_q^\delta$, uniformly for $\delta \in (0, 1]$, $x_0 \in \mathcal{K}$. In particular, any constant which is admissible (of any kind) in the sense of those results is ≈ 1 in the sense of this theorem (when working with ν , we only use 1; ν -admissible constants—see Definition 6.2 for the definition of 1; ν -admissible constants).

(vii) is contained in Theorem 4.7.

(ix): Theorem 4.7 gives $\xi_2 \approx 1$ ($\xi_2 < 1$) such that

$$B_{X^\delta}(x, \xi_2) \subseteq \Phi_{x,\delta}(B^n(\eta_1)) \subseteq B_{X^\delta}(x, 1) = B(x, \delta).$$

Thus, to prove (ix), we wish to show $\exists \xi_0 \approx 1$ with

$$B(x, \xi_0 \delta) \subseteq B_{X^\delta}(x, \xi_2). \quad (7.2)$$

Take $k \approx 1$ so large that $b_1^k \leq \xi_2$ and set $\xi_0 = B_1^{-k}$. Let $y \in B(x, \xi_0 \delta)$, so that there exists $\gamma : [0, 1] \rightarrow \Omega$, $\gamma(0) = x$, $\gamma(1) = y$, $\gamma'(t) = \langle a(t), X^{\xi_0 \delta}(\gamma(t)) \rangle$, with $\|a\|_{L^\infty} < 1$. Then,

$$\gamma'(t) = \left\langle a(t), \xi_2 A(t) X^\delta(\gamma(t)) \right\rangle = \left\langle A(t)^\top a(t), \xi_2 X^\delta(\gamma(t)) \right\rangle,$$

where $A(t) = \xi_2^{-1} S_{\xi_0 \delta}(\gamma(t)) S_{\xi_0 B_1 \delta}(\gamma(t)) \cdots S_{\xi_0 B_1^{k-1} \delta}(\gamma(t))$; note that $\|A\|_{L^\infty([0,1]; \mathbb{M}^{q \times q})} \leq 1$, and therefore, $\|A^\top a\|_{L^\infty([0,1]; \mathbb{R}^q)} \leq \|a\|_{L^\infty([0,1]; \mathbb{R}^q)} < 1$. It follows that $y = \gamma(1) \in B_{X^\delta}(x, \xi_2)$, completing the proof of (ix).

We claim, for $\delta_1 \leq \delta_2$,

$$\Lambda(x, \delta_1) \lesssim \Lambda(x, \delta_2), \quad (7.3)$$

where the implicit constant can be chosen to depend only on q . Indeed,

$$\begin{aligned} \Lambda(x, \delta_1) &= \max_{j_1, \dots, j_n \in \{1, \dots, q\}} \left| \det \left(X_{j_1}^{\delta_1}(x) | \cdots | X_{j_n}^{\delta_1}(x) \right) \right| \\ &= \max_{j_1, \dots, j_n \in \{1, \dots, q\}} \left| \det \left((T_{\delta_1, \delta_2} X^{\delta_2})_{j_1}(x) | \cdots | (T_{\delta_1, \delta_2} X^{\delta_2})_{j_n}(x) \right) \right|. \end{aligned}$$

Since $\|T_{\delta_1, \delta_2}(x)\| \leq 1$, the right hand side is the determinant of a matrix whose columns are linear combinations (with coefficients bounded by 1) of the vectors $X_1^{\delta_2}(x), \dots, X_q^{\delta_2}(x)$. (7.3) follows.

Next we claim, for $c > 0$ fixed,

$$\Lambda(x, c\delta) \approx \Lambda(x, \delta), \quad \delta, c\delta \in (0, 1], \quad (7.4)$$

where the implicit constant depends on c . It suffices to prove (7.4) for $c < 1$. By (7.3), it suffices to prove (7.4) for $c = B_2^{-k}$ for some k . We have

$$\begin{aligned} \Lambda(x, \delta) &= \max_{j_1, \dots, j_n \in \{1, \dots, q\}} \left| \det \left(X_{j_1}^\delta(x) | \cdots | X_{j_n}^\delta(x) \right) \right| \\ &= \max_{j_1, \dots, j_n \in \{1, \dots, q\}} \left| \det \left((AX^{c\delta})_{j_1}(x) | \cdots | (AX^{c\delta})_{j_n}(x) \right) \right|, \end{aligned} \quad (7.5)$$

where $A(x) = R_{B_2^{-1}\delta}(x) R_{B_2^{-2}\delta}(x) \cdots R_{B_2^{-k}\delta}(x)$. Since $\sup_{x \in \Omega} \|A(x)\|_{\mathbb{M}^{q \times q}} \leq b_2^{-k} \lesssim 1$, it follows that the right hand side of (7.5) is the determinant of a matrix whose columns are linear combinations (with coefficients whose magnitudes are $\lesssim 1$) of the vectors $X_1^{c\delta}(x), \dots, X_q^{c\delta}(x)$. It follows that $\Lambda(x, \delta) \lesssim \Lambda(x, c\delta)$. Combining this with (7.3), (7.4) follows.

Corollary 6.6 shows

$$\sigma_{\text{Leb}}(B_{X^\delta}(x, \xi_2)) \approx \Lambda(x, \delta), \quad (7.6)$$

where we have used that (thinking of σ_{Leb} as a density) $\sigma_{\text{Leb}}(V_1(x), \dots, V_n(x)) = |\det(V_1(x) \cdots V_n(x))|$. Combining this with (7.4) and (7.2), we have

$$\sigma_{\text{Leb}}(B(x, \xi_0 \delta)) \leq \sigma_{\text{Leb}}(B_{X^\delta}(x, \xi_2)) \approx \Lambda(x, \delta) \approx \Lambda(x, \xi_0 \delta). \quad (7.7)$$

Conversely, using (7.6) again we have,

$$\Lambda(x, \delta) \approx \sigma_{\text{Leb}}(B_{X^\delta}(x, \xi_2)) \leq \sigma_{\text{Leb}}(B_{X^\delta}(x, 1)) = \sigma_{\text{Leb}}(B(x, \delta)). \quad (7.8)$$

Combining (7.7) and (7.8) proves (v). (vi) follows from (v) and (7.4).

Since $\Phi_{x,\delta}^* \sigma_{\text{Leb}} = |\det d\Phi_{x,\delta}| \sigma_{\text{Leb}}$, (viii) follows from Theorem 6.5 (iii) and Corollary 6.6. (x) follows directly from Theorem 4.7. $\square$

REMARK 7.7. One can generalize the multi-parameter geometries from Section 7.2 in a similar way by changing the above variable $\delta \in (0, 1]$ to a *vector*, $\delta \in [0, 1]^\nu$ for some $\nu \in \mathbb{N}$, and proceeding in a similar way.

REMARK 7.8. The most artificial hypothesis in this section is (II). Indeed, it is not directly related to any of the hypotheses of a space of homogeneous type. This hypothesis can be replaced with weaker hypotheses and we can still achieve the same result. In fact, the main purposes of (II) are to ensure the existence of η and δ_0 (independent of $x \in \mathcal{K}$, $\delta \in (0, 1]$) in our application of Theorem 4.7, and to estimate $\mathcal{L}_{X_j^\delta} \sigma_{\text{Leb}}$. One could just directly assume the existence of such constants and estimates, or assume any number of other hypotheses which imply their existence, depending on the application at hand.

7.4 Diffeomorphism invariance and nonsmooth vector fields. An important way in which the results in this paper are stronger than previously mentioned works is that the statements of the main theorems are completely invariant under C^2 diffeomorphisms (see Section 3.2.1). This is true quantitatively: all of the estimates depend on quantities which are invariant under arbitrary C^2 diffeomorphisms. In previous works like [NSW85, TW03, Str11, MM12] the estimates were in terms of C^m type norms of the vector fields in some fixed coordinate system.¹¹ Thus, the vector fields had to be a priori “smooth” and “not large” in some fixed coordinate system. The concepts of “smooth” and “not large” are not invariant under C^2 diffeomorphisms. Under the assumptions of Theorem 4.7, we conclude the existence of a coordinate system in which the vector fields are smooth and not large, but we need not assume it. This allows us to address some settings where the vector fields are given in a coordinate system in which they are large and/or are merely C^1 ; in particular, unlike previous works, we only use the qualitative assumption that the vector fields are C^1 , and our estimates do not depend on the C^1 norms of the coefficients in a coordinate system.

¹¹ [MM12] works with Lipschitz vector fields to obtain some results with less regularity than the other mentioned works. It is possible that the ideas from that paper could be combined with the ideas from this paper to prove results like the ones in this paper, but with Lipschitz vector fields instead of C^1 vector fields; though we do not pursue this here.

When considering only the smoothness (and not the size) aspect of this diffeomorphism invariance, these results can be rephrased as the qualitative results in Section 3.1; the methods from previous works on this subject cannot yield such theorems, since they require the vector fields to be smooth in the first place. In fact, the qualitative results in this series seem to be of a new type; though there may be some connection to Hilbert’s fifth problem.

In the series of papers [SS11, Str12, SS13, SS12, Str17], the second author and Stein used the scaling techniques from [Str11] to study singular Radon transforms of the form

$$Tf(x) = \psi(x) \int f(\gamma(t, x))K(t) dt,$$

where $\gamma(t, x)$ is a germ of a smooth function defined near $(0, 0)$, $\gamma(t, x) : \mathbb{R}_0^N \times \mathbb{R}_0^n \rightarrow \mathbb{R}^n$ with $\gamma(0, x) \equiv x$ (we have used $\mathbb{R}_0^m$ to denote a small neighborhood of $0 \in \mathbb{R}^m$), and $K(t)$ is a “multi-parameter singular kernel” supported near $0 \in \mathbb{R}^N$. Conditions were given so that the above operator was bounded on L^p . Because the theory was based on [Str11], it was required that $\gamma(t, x)$ be smooth and supported very near $(0, 0)$. One could replace every application of the results from [Str11] in these papers with Theorem 4.7 to obtain more general results where γ is not necessarily required to be smooth or supported very close to 0. In fact, the results can be made completely invariant under arbitrary C^2 diffeomorphisms, and so the concepts of smooth and small do not have intrinsic meaning. Similar remarks hold for many other settings where methods from [NSW85, TW03, Str11] are used.

Large sub-Riemannian balls have been studied in some special cases before. See, for example, the discussion of model pseudoconvex boundaries in [NS01, Section 4] as well as [Pet14, DP18]. The approach in this paper allows us to unify the ideas behind these large sub-Riemannian balls with the more robust theory of small sub-Riemannian balls.

7.5 Several complex variables. As described in Section 7.3, the results in this paper can be used to study generalized versions of sub-Riemannian geometries, and as elucidated by Charpentier and Dupain [CD14], these geometries arise when studying $\bar{\partial}$ -problems. When applying the results from this series to such questions, a difficulty arises. We turn to describing this issue, and how it will be addressed in a future work of the second author.

Let M be a complex manifold of dimension n , and for each $\delta \in (0, 1]$, let $L_1^\delta, \dots, L_q^\delta$ be C^1 complex vector fields on M such that $\forall \zeta \in M$, $\text{span}\{L_1^\delta(\zeta), \dots, L_q^\delta(\zeta)\} = T_\zeta^{0,1}M$. Let $X_1^\delta, \dots, X_{2q}^\delta$ denote the list of real vector fields $\text{Re}(L_1^\delta), \dots, \text{Re}(L_q^\delta), \text{Im}(L_1^\delta), \dots, \text{Im}(L_q^\delta)$. We assume that the list $X_1^\delta, \dots, X_{2q}^\delta$ locally satisfies the hypotheses of Section 7.3. Then, Theorem 7.6 applies to show that the balls $B(x, \delta)$ defined in that section locally give M the structure of a space of homogeneous type¹²,

¹² Since M is an abstract manifold, we do not have a natural choice of density σ_{Leb} on M . However, one may instead use any strictly positive C^1 density on M and obtain the same results. All such choices of density are equivalent for our purposes.

and we obtain scaling maps $\Phi_{x,\delta} : B^{2n}(\eta_1) \rightarrow B(x, \delta)$ as in that theorem. In particular, by Theorem 7.6 (x), the maps $\Phi_{x,\delta}$ “rescale” the vector fields $X_1^\delta, \dots, X_{2q}^\delta$ so that they are smooth and span the tangent space, uniformly for x in compact sets and $\delta \in (0, 1]$.

In other words, $\Phi_{x,\delta}^* L_1^\delta, \dots, \Phi_{x,\delta}^* L_q^\delta, \Phi_{x,\delta}^* \overline{L_1^\delta}, \dots, \Phi_{x,\delta}^* \overline{L_q^\delta}$ are smooth and span the complexified tangent space, uniformly for x in compact sets and $\delta \in (0, 1]$. The hope is to apply techniques from several complex variables at the unit scale to these rescaled vector fields, to be able to conclude results at every scale $\delta \in (0, 1]$. However, there is one key component that is missing in the complex setting. We identify $\mathbb{R}^{2n}$ with $\mathbb{C}^n$ via the map $(x_1, \dots, x_{2n}) \mapsto (x_1 + ix_{n+1}, \dots, x_n + ix_{2n})$. To be able to apply results from complex analysis, we would need that $\Phi_{x,\delta}^* L_1^\delta, \dots, \Phi_{x,\delta}^* L_q^\delta$ (thought of as vector fields on the ball of radius η_1 in $\mathbb{C}^n$) are still $T^{0,1}$ vector fields. It is easy to see that this is equivalent to the map $\Phi_{x,\delta}$ being holomorphic. However, the best one can say about the maps constructed in this series is that they are C^2 .

One therefore wishes to obtain the same results as this paper, but with a different map Φ , where we can also conclude that Φ is holomorphic. In the past, this has been achieved in special cases by using ad hoc methods for the particular problem at hand (e.g., by using non-isotropic dilations determined by the Taylor series of some ingredients in the problem)—see, for example, [NRSW89, Section 3], [CD14, Section 3.3.2], and [CD06, Section 2.1]. However, using such ad hoc methods does not allow one to proceed in the generality of this paper, and can obfuscate the underlying mechanism of the problem.

In a forthcoming paper, the second author will address this issue, and obtain appropriate analogs of results in this series in the complex setting; which can be seen as a quantitatively diffeomorphic invariant version of the classical Newlander–Nirenberg theorem [NN57]. The results and methods of this series are the first step in addressing this complex setting.

When we move to the complex setting (and more general settings which will be discussed in a future paper), the ODE methods of this paper are no longer sufficient to obtain even non-sharp results, and one must move to PDE methods. In particular, Zygmund spaces are the right scale of spaces to discuss any of the results in the complex setting.

8 Function Spaces, Revisited

In this section, we state and prove the basic results we need concerning the function spaces introduced in Section 2. We begin with several straightforward inclusions of these spaces, which we state in the next lemma. For the rest of this section, we take the setting of Section 2.2.

LEMMA 8.1. (i) For $0 \leq s_1 \leq s_2 \leq 1$, $m \in \mathbb{N}$, $\|f\|_{C_X^{m,s_1}(M)} \leq 3\|f\|_{C_X^{m,s_2}(M)}$.
(ii) $\|f\|_{C_X^{m,1}(M)} \leq \|f\|_{C_X^{m+1}(M)}$.
(iii) For $s \in (0, 1]$, $m \in \mathbb{N}$, $\|f\|_{C_X^{s+m}(M)} \leq 5\|f\|_{C_X^{m,s}(M)}$.

- (iv) For $0 < s_1 \leq s_2 < \infty$, $\|f\|_{\mathcal{C}_X^{s_1}(M)} \leq 15\|f\|_{\mathcal{C}_X^{s_2}(M)}$.
(v) If $U \subseteq M$ is an open set, then $\|f\|_{C_X^{m,s}(U)} \leq \|f\|_{C_X^{m,s}(M)}$ and $\|f\|_{\mathcal{C}_X^s(U)} \leq \|f\|_{\mathcal{C}_X^s(M)}$.

Proof. For (i), it suffices to prove the case $m = 0$. We have,

$$\begin{aligned} \|f\|_{C_X^{0,s_1}} &= \|f\|_{C(M)} + \sup_{x \neq y} \rho(x, y)^{-s_1} |f(x) - f(y)| \\ &\leq \|f\|_{C(M)} + \sup_{x \neq y} \min\{\rho(x, y), 1\}^{-s_1} |f(x) - f(y)| \\ &\leq \|f\|_{C(M)} + \sup_{x \neq y} \min\{\rho(x, y), 1\}^{-s_2} |f(x) - f(y)| \\ &\leq 3\|f\|_{C(M)} + \sup_{x \neq y} \rho(x, y)^{-s_2} |f(x) - f(y)| \leq 3\|f\|_{C_X^{0,s_2}}, \end{aligned}$$

proving (i).

For (ii), it suffices to prove the case $m = 0$. Let $x \neq y \in M$ with $\rho(x, y) < \infty$, fix $\epsilon > 0$, and let $\delta = \rho(x, y) + \epsilon$. Pick $\gamma : [0, 1] \rightarrow M$ with $\gamma(0) = x$, $\gamma(1) = y$, $\gamma'(t) = \sum_{j=1}^q a_j(t) \delta X_j(\gamma(t))$, $\|\sum |a_j|^2\|_{L^\infty([0,1])} < 1$. Then we have,

$$\begin{aligned} &\rho(x, y)^{-1} |f(x) - f(y)| \\ &= \rho(x, y)^{-1} \left| \int_0^1 \sum a_j(t) \delta(X_j f)(\gamma(t)) dt \right| \\ &\leq \frac{\delta}{\rho(x, y)} \left\| \max_{1 \leq j \leq q} |a_j(t)| \right\|_{L^\infty([0,1])} \sum_{j=1}^q \|X_j f\|_{C(M)} \\ &\leq \frac{\delta}{\rho(x, y)} \sum_{j=1}^q \|X_j f\|_{C(M)} = \frac{\rho(x, y) + \epsilon}{\rho(x, y)} \sum_{j=1}^q \|X_j f\|_{C(M)} \xrightarrow{\epsilon \rightarrow 0} \sum_{j=1}^q \|X_j f\|_{C(M)}. \end{aligned}$$

If $\rho(x, y) = \infty$, then $\rho(x, y)^{-1} |f(x) - f(y)| = 0 \leq \sum_j \|X_j f\|_{C(M)}$. It follows that $\|f\|_{C_X^{0,1}(M)} \leq \|f\|_{C_X^1(M)}$, completing the proof of (ii).

For (iii), it suffices to prove the case $m = 0$. Let $\gamma \in \mathcal{P}_{X,s/2}^M(h)$. Then $\rho(\gamma(2h), \gamma(h)), \rho(\gamma(h), \gamma(0)) < h$, and so

$$\begin{aligned} h^{-s} |f(\gamma(2h)) - 2f(\gamma(h)) + f(\gamma(0))| &\leq 2 \sup_{\rho(x,y) < h} h^{-s} |f(x) - f(y)| \\ &\leq 2 \sup_{x,y \in M, x \neq y} \rho(x, y)^{-s} |f(x) - f(y)|. \end{aligned}$$

Combining this with $\|f\|_{C_X^{0,s/2}(M)} \leq 3\|f\|_{C_X^{0,s}(M)}$ (by (i)), (iii) follows.

For (iv) it suffices to prove the case when $s_1 \in (0, 1]$. When $s_2 \in (0, 1]$, as well, then it follows easily from the definitions that $\|f\|_{\mathcal{C}_X^{s_1}(M)} \leq 5\|f\|_{\mathcal{C}_X^{s_2}(M)}$. When $s_2 > 1$, we use (iii), (i), and (ii) to see

$$\|f\|_{\mathcal{C}_X^{s_1}(M)} \leq 5\|f\|_{C_X^{0,s_1}(M)} \leq 15\|f\|_{C_X^{0,1}(M)} \leq 15\|f\|_{C_X^1(M)} \leq 15\|f\|_{\mathcal{C}_X^{s_2}(M)},$$

completing the proof of (iv). (v) follows easily from the definitions. $\square$

REMARK 8.2. Given the analogy with Euclidean spaces, one expects the reverse inequality to Lemma 8.1 (iii), when $s \in (0, 1)$; namely $\|f\|_{C_X^{m,s}(M)} \lesssim \|f\|_{\mathcal{C}_X^{s+m}(M)}$. Under additional hypotheses, this is true locally. See the second paper in this series for details.

PROPOSITION 8.3. *The spaces $C_X^{m,s}(M)$, $\mathcal{C}_X^s(M)$, $C^{m,s}(\Omega)$, and $\mathcal{C}^s(\Omega)$ are algebras. In fact, we have for $m \in \mathbb{N}$, $s \in [0, 1]$,*

$$\|fg\|_{C_X^{m,s}(M)} \leq C_{m,q} \|f\|_{C_X^{m,s}(M)} \|g\|_{C_X^{m,s}(M)},$$

where $C_{m,q}$ is a constant depending only on m and q . And for $m \in \mathbb{N}$, $s \in (m, m+1]$,

$$\|fg\|_{\mathcal{C}_X^s(M)} \leq C_{m,q} \|f\|_{\mathcal{C}_X^s(M)} \|g\|_{\mathcal{C}_X^s(M)}. \quad (8.1)$$

Moreover, these algebras have multiplicative inverses for functions which are bounded away from zero. If $f \in C_X^{m,s}(M)$ with $\inf_{x \in M} |f(x)| \geq c_0 > 0$ then $f(x)^{-1} = \frac{1}{f(x)} \in C_X^{m,s}(M)$ with

$$\|f(x)^{-1}\|_{C_X^{m,s}(M)} \leq C,$$

where C can be chosen to depend only on m , q , c_0 , and an upper bound for $\|f\|_{C_X^{m,s}(M)}$. And for $m \in \mathbb{N}$, $s \in (m, m+1]$ if $f \in \mathcal{C}_X^s(M)$ with $\inf_{x \in M} |f(x)| \geq c_0 > 0$ then $f(x)^{-1} \in \mathcal{C}_X^s(M)$ with

$$\|f(x)^{-1}\|_{\mathcal{C}_X^s(M)} \leq C, \quad (8.2)$$

where C can be chosen to depend only on m , q , c_0 , and an upper bound for $\|f\|_{\mathcal{C}_X^s(M)}$. The same results hold with $C_X^{m,s}(M)$ replaced by $C^{m,s}(\Omega)$ and $\mathcal{C}_X^s(M)$ replaced by $\mathcal{C}^s(\Omega)$ (with n playing the role of q).

Proof. The proofs for $C_X^{m,s}(M)$ and $C^{m,s}(\Omega)$ are straightforward and standard, so we focus on the Zygmund spaces. We prove (8.1) by induction on m , where $s \in (m, m+1]$. We begin with the base case $s \in (0, 1]$. Since we already know $\|fg\|_{C_X^{0,s/2}(M)} \lesssim \|f\|_{C_X^{0,s/2}(M)} \|g\|_{C_X^{0,s/2}(M)}$, it suffices to show for $\gamma \in \mathcal{P}_{X,s/2}^M(h)$,

$$h^{-s} |f(\gamma(2h))g(\gamma(2h)) - 2f(\gamma(h))g(\gamma(h)) + f(\gamma(0))g(\gamma(0))| \leq 6 \|f\|_{\mathcal{C}_X^s(M)} \|g\|_{\mathcal{C}_X^s(M)}.$$

Notice that $\rho(\gamma(h), \gamma(0)) \leq h$, and therefore $|f(\gamma(h)) - f(\gamma(0))| \leq h^{s/2} \|f\|_{C_X^{0,s/2}(M)}$. Thus, we have

$$\begin{aligned} & h^{-s} |f(\gamma(2h))g(\gamma(2h)) - 2f(\gamma(h))g(\gamma(h)) + f(\gamma(0))g(\gamma(0))| \\ & \leq h^{-s} |f(\gamma(2h)) - 2f(\gamma(h)) + f(\gamma(0))| |g(\gamma(2h))| \\ & \quad + h^{-s} |2f(\gamma(h)) - f(\gamma(0))| |g(\gamma(2h)) - 2g(\gamma(h)) + g(\gamma(0))| \\ & \quad + h^{-s} 2 |f(\gamma(h)) - f(\gamma(0))| |g(\gamma(h)) - g(\gamma(0))| \\ & \leq \|f\|_{\mathcal{C}_X^s(M)} \|g\|_{C(M)} + 3 \|f\|_{C(M)} \|g\|_{\mathcal{C}_X^s(M)} + 2 \|f\|_{C_X^{0,s/2}(M)} \|g\|_{C_X^{0,s/2}(M)} \\ & \leq 6 \|f\|_{\mathcal{C}_X^s(M)} \|g\|_{\mathcal{C}_X^s(M)}. \end{aligned}$$

Having proved the base case, (8.1) follows by a straightforward induction, which we leave to the reader.

We now turn to inverses. We prove (8.2) by induction on m , where $s \in (m, m+1]$. We begin with the base case $s \in (0, 1]$. Let $f \in \mathcal{C}_X^s(M)$ with $\inf_{x \in M} |f(x)| \geq c_0 > 0$. We write $A \lesssim B$ for $A \leq CB$ where C is as in (8.2). Since we already know the results for Hölder spaces, we have $\|f(x)^{-1}\|_{C_X^{0,s/2}(M)} \lesssim 1$. Thus, it suffices to show for $\gamma \in \mathcal{P}_{X,s/2}^M(h)$,

$$\begin{aligned} & \left| \frac{1}{f(\gamma(2h))} - \frac{2}{f(\gamma(h))} + \frac{1}{f(\gamma(0))} \right| \\ &= \left| \frac{f(\gamma(h))f(\gamma(0)) - 2f(\gamma(2h))f(\gamma(0)) + f(\gamma(2h))f(\gamma(h))}{f(\gamma(2h))f(\gamma(h))f(\gamma(0))} \right| \lesssim h^s. \end{aligned}$$

Since we have $|f(\gamma(2h))f(\gamma(h))f(\gamma(0))| \geq c_0^3 \gtrsim 1$, it suffices to show

$$|f(\gamma(h))f(\gamma(0)) - 2f(\gamma(2h))f(\gamma(0)) + f(\gamma(2h))f(\gamma(h))| \lesssim h^s.$$

But we have

$$\begin{aligned} & |f(\gamma(h))f(\gamma(0)) - 2f(\gamma(2h))f(\gamma(0)) + f(\gamma(2h))f(\gamma(h))| \\ & \leq |(f(\gamma(2h)) - 2f(\gamma(h)) + f(\gamma(0)))f(\gamma(h))| + 2|f(\gamma(h))^2 - f(\gamma(2h))f(\gamma(0))| \\ & \leq h^s \|f\|_{\mathcal{C}_X^s(M)} \|f\|_{C(M)} + 2|f(\gamma(h))^2 - f(\gamma(2h))f(\gamma(0))| \\ & \lesssim h^s + 2|f(\gamma(h))^2 - f(\gamma(2h))f(\gamma(0))|. \end{aligned}$$

Thus, it suffices to show

$$|f(\gamma(h))^2 - f(\gamma(2h))f(\gamma(0))| \lesssim h^s.$$

But, using that $\rho(\gamma(h), \gamma(0)) \leq h$, and therefore $|f(\gamma(h)) - f(\gamma(0))| \leq h^{s/2} \|f\|_{C_X^{0,s/2}(M)} \lesssim h^{s/2}$, we have

$$\begin{aligned} & |f(\gamma(h))^2 - f(\gamma(2h))f(\gamma(0))| \\ & \leq |(f(\gamma(2h)) - 2f(\gamma(h)) + f(\gamma(0)))f(\gamma(0))| + |(f(\gamma(h)) - f(\gamma(0)))^2| \\ & \lesssim h^s + h^s \lesssim h^s, \end{aligned}$$

completing the proof of the base case. Having proved the base case, the inductive step is straightforward, and we leave it to the reader.

The proofs for $\mathcal{C}^s(\Omega)$ are similar, and we leave them to the reader. $\square$

REMARK 8.4. In the proof of Proposition 8.3, it is used that $\|f\|_{C^{0,s/2}(\Omega)} \leq \|f\|_{\mathcal{C}^s(\Omega)}$, $s \in (0, 1]$, which is clearly true because of our nonstandard definition of $\|f\|_{\mathcal{C}^s(\Omega)}$ (see Remark 2.1). Even with the more standard definition, for a bounded Lipschitz domain Ω , one has $\|f\|_{C^{0,s/2}(\Omega)} \leq C\|f\|_{\mathcal{C}^s(\Omega)}$, however C depends on Ω . Thus, if one takes the more standard definition, the conclusions of Proposition 8.3 take a more complicated form.

REMARK 8.5. Lemma 8.1 and Proposition 8.3 hold (with exactly the same proofs) if M is replaced by $B_X(x_0, \xi)$, whether or not $B_X(x_0, \xi)$ is a manifold—see Section 2.2.1.

PROPOSITION 8.6. *Let N be another C^2 manifold, $Y_1, \dots, Y_q$ be C^1 vector fields on N , and $\Phi : N \rightarrow M$ be a C^1 map such that $d\Phi(u)Y_j(u) = X_j(\Phi(u))$, $\forall u \in N$. Then,*

$$\|f \circ \Phi\|_{C_Y^{m,s}(N)} \leq \|f\|_{C_X^{m,s}(M)}, \quad m \in \mathbb{N}, s \in [0, 1], \quad (8.3)$$

$$\|f \circ \Phi\|_{\mathcal{C}_Y^s(N)} \leq \|f\|_{\mathcal{C}_X^s(M)}, \quad s > 0. \quad (8.4)$$

Proof. We begin with (8.3). Since $Y^\alpha(f \circ \Phi) = (X^\alpha f) \circ \Phi$, it suffices to prove the case $m = 0$. We have a sub-Riemannian metric ρ_Y on N and another sub-Riemannian metric ρ_X on M , defined by (2.3). We claim

$$\rho_X(\Phi(u_1), \Phi(u_2)) \leq \rho_Y(u_1, u_2). \quad (8.5)$$

This is clear if $\rho_Y(u_1, u_2) = \infty$. If $\rho_Y(u_1, u_2) < \infty$, let $\delta > \rho_Y(u_1, u_2)$. Then, there exists $\gamma : [0, 1] \rightarrow N$, $\gamma(0) = u_1$, $\gamma(1) = u_2$, $\gamma'(t) = \sum a_j(t)\delta Y_j(\gamma(t))$, $\|\sum |a_j|^2\|_{L^\infty([0,1])} < 1$. Set $\tilde{\gamma} = \Phi \circ \gamma$. Then, $\tilde{\gamma}(0) = \Phi(u_1)$, $\tilde{\gamma}(1) = \Phi(u_2)$, and $\tilde{\gamma}'(t) = \sum a_j(t)\delta X_j(\tilde{\gamma}(t))$. This proves $\rho_X(\Phi(u_1), \Phi(u_2)) < \delta$. Taking $\delta \rightarrow \rho_Y(u_1, u_2)$ proves (8.5). We conclude, for $s \in [0, 1]$,

$$\rho_Y(u_1, u_2)^{-s} |f \circ \Phi(u_1) - f \circ \Phi(u_2)| \leq \rho_X(\Phi(u_1), \Phi(u_2))^{-s} |f(\Phi(u_1)) - f(\Phi(u_2))|.$$

(8.3) follows.

We turn to (8.4). Again, since $Y^\alpha(f \circ \Phi) = (X^\alpha f) \circ \Phi$, it suffices to prove (8.4) for $s \in (0, 1]$. That $\|f \circ \Phi\|_{C_Y^{0,s/2}(N)} \leq \|f\|_{C_X^{0,s/2}(M)}$ follows from (8.3). Furthermore, it follows easily from the definitions that for $\gamma \in \mathcal{P}_{Y,s/2}^N(h)$, we have $\Phi \circ \gamma \in \mathcal{P}_{X,s/2}^M(h)$. Using this, (8.4) for $s \in (0, 1]$ follows immediately. $\square$

8.1 Comparison with Euclidean function spaces. Fix $\eta \in (0, 1]$ and let $Y_1, \dots, Y_q$ be vector fields on $B^n(\eta)$. When $Y_1, \dots, Y_q$ span the tangent space at every point of $B^n(\eta)$ and are sufficiently smooth, we have $C_Y^{m,s}(B^n(\eta)) = C^{m,s}(B^n(\eta))$ and $\mathcal{C}_Y^s(B^n(\eta)) = \mathcal{C}^s(B^n(\eta))$. In what follows, we state and prove quantitative versions of these equalities.

We write $Y_j = \sum_{k=1}^n a_j^k \frac{\partial}{\partial t_k}$ and assume $\frac{\partial}{\partial t_k} = \sum_{j=1}^q b_k^j Y_j$, where $a_j^k \in C^1(B^n(\eta))$, $b_k^j \in C(B^n(\eta))$.

DEFINITION 8.7. *In analogy with Definition 2.2, for $m < 0$ we define $C^{m,s}(B^n(\eta)) := C(B^n(\eta))$, with equality of norms. For $s \in (-1, 0]$ we define $\mathcal{C}^s(B^n(\eta)) := C^{0,(s+1)/2}(B^n(\eta))$, with equality of norms.*

DEFINITION 8.8. *We say C is a 0:E-admissible constant¹³ if C can be chosen to depend only on upper bounds for q and $\|a_j^k\|_{C(B^n(\eta))}$, $\|b_k^j\|_{C(B^n(\eta))}$, $\forall j, k$.*

¹³ Here we are using the E to stand for Euclidean, and to help differentiate these admissible constants from the other admissible constants in this paper.

DEFINITION 8.9. For $m \in \mathbb{Z}$, $s \in [0, 1]$, if we say C is an $\langle m, s; E \rangle$ -admissible constant if $a_j^k, b_k^j \in C^{m,s}(B^n(\eta))$, $\forall j, k$, and C can be chosen to depend only on upper bounds for q , m , and $\|a_j^k\|_{C^{m,s}(B^n(\eta))}$, $\|b_k^j\|_{C^{m,s}(B^n(\eta))}$, $\forall j, k$.

DEFINITION 8.10. For $s > -1$ we say C is an $\{s; E\}$ -admissible constant if $a_j^k, b_k^j \in \mathcal{C}^s(B^n(\eta))$, $\forall j, k$ and C can be chosen to depend only on s and upper bounds for q , η^{-1} , and $\|a_j^k\|_{\mathcal{C}^s(B^n(\eta))}$, $\|b_k^j\|_{\mathcal{C}^s(B^n(\eta))}$, $\forall j, k$.

As before, we define $A \lesssim_{\langle m, s; E \rangle} B$ to be $A \leq CB$ where C is an $\langle m, s; E \rangle$ -admissible constant. We similarly define $\approx_{\langle m, s; E \rangle}$, $\lesssim_{\{s; E\}}$, and $\approx_{\{s; E\}}$. Recall, the vector fields $Y_1, \dots, Y_q$ induce a metric ρ on $B^n(\eta)$ via (2.3).

LEMMA 8.11. $\rho(x, y) \approx_{0; E} |x - y|$.

Proof. This follows immediately from the assumptions. $\square$

PROPOSITION 8.12. For $m \in \mathbb{N}$, $s \in [0, 1]$,

$$\|f\|_{C^{m,s}(B^n(\eta))} \approx_{\langle m-1, s; E \rangle} \|f\|_{C_Y^{m,s}(B^n(\eta))}, \quad (8.6)$$

and for $s > 0$,

$$\|f\|_{\mathcal{C}^s(B^n(\eta))} \approx_{\{s-1; E\}} \|f\|_{\mathcal{C}_Y^s(B^n(\eta))}. \quad (8.7)$$

Proof. We use Proposition 8.3 freely in this proof. In this proof, the norms $\|f\|_{C_Y^{m,s}}$, $\|f\|_{C^{m,s}}$, $\|f\|_{\mathcal{C}_Y^s}$, and $\|f\|_{\mathcal{C}^s}$ are always taken to be over the domain $B^n(\eta)$ unless otherwise mentioned. We prove (8.6) by induction on m . The base case, $\|f\|_{C^{0,s}(B^n(\eta))} \approx_{0; E} \|f\|_{C_Y^{0,s}(B^n(\eta))}$, follows immediately from Lemma 8.11. We assume (8.6) for $m-1$ and prove it for m . We have

$$\begin{aligned} \|f\|_{C_Y^{m,s}} &= \|f\|_{C_Y^{m-1,s}} + \sum_{j=1}^q \|Y_j f\|_{C_Y^{m-1,s}} \approx_{\langle m-2, s; E \rangle} \|f\|_{C^{m-1,s}} + \sum_{j=1}^q \|Y_j f\|_{C^{m-1,s}} \\ &\leq \|f\|_{C^{m-1,s}} + \sum_{j=1}^q \sum_{k=1}^n \|a_j^k \partial_{x_k} f\|_{C^{m-1,s}} \lesssim_{\langle m-1, s; E \rangle} \|f\|_{C^{m,s}}. \end{aligned}$$

For the reverse inequality,

$$\begin{aligned} \|f\|_{C^{m,s}} &\leq \|f\|_{C^{m-1,s}} + \sum_{k=1}^n \|\partial_{x_k} f\|_{C^{m-1,s}} \leq \|f\|_{C^{m-1,s}} + \sum_{k=1}^n \sum_{j=1}^q \|b_k^j Y_j f\|_{C^{m-1,s}} \\ &\lesssim_{\langle m-1, s; E \rangle} \|f\|_{C^{m-1,s}} + \sum_{j=1}^q \|Y_j f\|_{C^{m-1,s}} \lesssim_{\langle m-2, s; E \rangle} \|f\|_{C_Y^{m-1,s}} + \sum_{j=1}^q \|Y_j f\|_{C_Y^{m-1,s}} \\ &= \|f\|_{C_Y^{m,s}}. \end{aligned}$$

This completes the proof of (8.6).

We prove (8.7) by induction on m , where $s \in (m, m+1]$. We begin with the base case, $m = 0$, and thus $s \in (0, 1]$. First we show $\lesssim_{\{s-1; E\}}$. Take $0 \neq h \in \mathbb{R}^n$, and $x \in \Omega_h$ (where $\Omega = B^n(\eta)$). Set $\gamma(t) = x + t\theta$, where $\theta = h/|h|$. Note $\gamma'(t) = \sum_{k=1}^n \theta_k \frac{\partial}{\partial x_k} = \sum_{j=1}^q \sum_{k=1}^n \theta_k b_k^j(\gamma(t)) Y_j(\gamma(t))$. Since $\|b_k^j\|_{C^{0,s/2}} \lesssim_{\{s-1; E\}} 1$, we have $\|b_j^k \circ \gamma\|_{C^{0,s/2}} \lesssim_{\{s-1; E\}} 1$, and therefore $\gamma \in \mathcal{P}_{Y,s/2}^{B^n(\eta)}(C|h|)$, where $C \lesssim_{\{s-1; E\}} 1$. Hence,

$$\begin{aligned} & |h|^{-s} |f(x+2h) - 2f(x+h) + f(x)| \\ & \lesssim_{\{s-1; E\}} (C|h|)^{-s} |f(\gamma(2|h|)) - 2f(\gamma(|h|)) + f(\gamma(0))| \\ & \leq \|f\|_{\mathcal{C}_Y^s(B^n(\eta))}. \end{aligned}$$

Since we have already shown $\|f\|_{C^{0,s/2}} \approx_{\{s-1; E\}} \|f\|_{C_Y^{0,s/2}}$ (by (8.6)), the $\lesssim_{\{s-1; E\}}$ direction of (8.7) follows.

We turn to $\gtrsim_{\{s-1; E\}}$. We already have $\|f\|_{C_Y^{0,s/2}} \approx_{\{s-1; E\}} \|f\|_{C^{0,s/2}} \leq \|f\|_{\mathcal{C}^s}$ (by (8.6)). Fix $h > 0$ and $\gamma \in \mathcal{P}_{Y,s/2}^{B^n(\eta)}(h)$. Note $\gamma'(t) = \sum_{j=1}^q d_j(t) Y_j(\gamma(t)) = \sum_{j=1}^q \sum_{k=1}^n d_j(t) a_j^k(\gamma(t)) \frac{\partial}{\partial x_k}$, with $\sum \|d_j\|_{C^{0,s/2}([0,2h])}^2 < 1$. Since we also have $\|a_j^k\|_{C^{0,s/2}} \lesssim_{\{s-1; E\}} 1$, it follows that $\|\gamma\|_{C^{1,s/2}([0,2h])} \lesssim_{\{s-1; E\}} 1$. Define $\tilde{\gamma} : [0, 2h] \rightarrow B^n(\eta)$ by $\tilde{\gamma}(t) = (t/2h)\gamma(2h) + (1 - t/2h)\gamma(0)$.

We claim that

$$|\gamma(t) - \tilde{\gamma}(t)| \lesssim_{\{s-1; E\}} h^{1+s/2}. \quad (8.8)$$

Indeed,

$$|\tilde{\gamma}(t) - \gamma(t)| = t \left| \frac{\gamma(2h) - \gamma(0)}{2h} - \frac{\gamma(t) - \gamma(0)}{t} \right| = t |\gamma'(c_1) - \gamma'(c_2)|,$$

by the mean value theorem, where $c_1, c_2 \in [0, 2h]$. Since $t \in [0, 2h]$, it follows that $|\gamma(t) - \tilde{\gamma}(t)| \lesssim_{\{s-1; E\}} h^{1+s/2}$, by using the estimate $\|\gamma\|_{C^{1,s/2}([0,2h])} \lesssim_{\{s-1; E\}} 1$.

Next we claim that

$$\|f\|_{C^{0,s/(1+s/2)}(B^n(\eta))} \lesssim_{\{s-1; E\}} \|f\|_{\mathcal{C}^s(B^n(\eta))}. \quad (8.9)$$

To prove (8.9) we use

$$\|f\|_{C^{0,s/(1+s/2)}(B^n(\eta))} \approx \|f\|_{\mathcal{C}^{s/(1+s/2)}(B^n(\eta))}, \quad (8.10)$$

where the implicit constants depend on s, n , and an upper bound for η^{-1} (here we use $s/(1+s/2) \in (0, 1)$; (8.10) does not hold when the exponent equals 1). Then, since $0 < s/(1+s/2) < s \leq 1$, we have $\|f\|_{\mathcal{C}^{s/(1+s/2)}(B^n(\eta))} \leq 5\|f\|_{\mathcal{C}^s(B^n(\eta))}$ (this follows immediately from the definitions) and (8.9) follows. (8.10) is classical; indeed, we first consider the case when $\eta = 1$. The $\gtrsim$ part of (8.10) follows immediately from the definitions. For the $\lesssim$ part when $\eta = 1$, see [Tri06, Theorem 1.118 (i)]—by choosing $M = 1, 2$ in that theorem, the $\lesssim$ part of (8.10) follows, for $\eta = 1$, with implicit constant depending only on s and n . Finally, a simple scaling argument establishes (8.10) for general $\eta > 0$, which we leave to the reader.

Note that $\tilde{\gamma}(t)$ is a line with $|\tilde{\gamma}(2h) - \tilde{\gamma}(0)| \leq 2h\|\gamma\|_{C^1([0,2h])} \lesssim_{\{s-1:E\}} h$; and therefore $|f(\tilde{\gamma}(2h)) - 2f(\tilde{\gamma}(h)) + f(\tilde{\gamma}(0))| \lesssim_{\{s-1:E\}} h^s \|f\|_{\mathcal{C}^s}$. We combine this with (8.8) and (8.9) to see:

$$\begin{aligned} & |f(\gamma(2h)) - 2f(\gamma(h)) + f(\gamma(0))| \\ & \leq |f(\tilde{\gamma}(2h)) - 2f(\tilde{\gamma}(h)) + f(\tilde{\gamma}(0))| + 2|f(\gamma(h)) - f(\tilde{\gamma}(h))| \\ & \lesssim_{\{s-1:E\}} h^s \|f\|_{\mathcal{C}^s} + |\gamma(h) - \tilde{\gamma}(h)|^{s/(1+s/2)} \|f\|_{C^{0,s/(1+s/2)}} \\ & \lesssim_{\{s-1:E\}} h^s \|f\|_{\mathcal{C}^s}. \end{aligned}$$

This proves $\|f\|_{\mathcal{C}^s_\gamma} \lesssim_{\{s-1:E\}} \|f\|_{\mathcal{C}^s}$, and completes the proof for the base case of (8.7). From here the inductive step follows just as in the inductive step for (8.6), and we leave it to the reader. $\square$

9 Proofs

We turn to the proofs of the main results of this paper. The heart of this paper is the study of a certain ODE which arises in canonical coordinates; this is presented in Section 9.1. Then we present a quantitative version of a special case of the Inverse Function Theorem in Section 9.2. We then prove the main result (Theorem 4.7) in Section 9.3. Next, we prove the results concerning densities from Section 6 in Section 9.4. Finally, we prove Proposition 4.14 in Section 9.5.

9.1 An ODE. The quantitative study of canonical coordinates is closely tied to the study of the following ODE, defined for an $n \times n$ matrix $A(u)$, depending on $u \in B^n(\eta)$ for some $\eta > 0$. Write $u = r\theta$, $r > 0$, $\theta \in S^{n-1}$. The ODE is:

$$\frac{\partial}{\partial r} rA(r\theta) = -A(r\theta)^2 - C(r\theta)A(r\theta) - C(r\theta), \quad (9.1)$$

where $C(u) \in C(B^n(\eta); \mathbb{M}^{n \times n})$ is a given function. That this ODE arises in the study of canonical coordinates is classical (see, for example, [Che46, p. 155] for the derivation of a similar ODE); however the detailed study of the ODE to prove regularity properties in canonical coordinates was pioneered by Tao and Wright [TW03].

In Section 9.1.1 we show how this ODE arises in canonical coordinates. Because our vector fields $X_1, \dots, X_q$ are merely assumed to be C^1 , there are some slight technicalities which we deal with in that section. In Section 9.1.2 we prove the regularity properties of solutions to this ODE.

9.1.1 Derivation of the ODE. Let $X_1, \dots, X_n$ be C^1 vector fields on an n -dimensional C^2 manifold M . Fix $x \in M$ and $\epsilon > 0$ and suppose:

- $X_1, \dots, X_n$ span the tangent space at every point of M .
- $\Phi(u) := e^{u_1 X_1 + u_2 X_2 + \dots + u_n X_n} x$ exists for $u \in B^n(\epsilon)$.

Write $[X_j, X_k] = \sum_{l=1}^n c_{j,k}^l X_l$. Since $X_1, \dots, X_n$ form a basis for the tangent space of M at every point, $c_{j,k}^l \in C(M)$ are uniquely defined. Classical theorems show that Φ is C^1 (since $X_1, \dots, X_n$ are).

Let $U \subseteq M$ and $V \subseteq B^n(\epsilon)$ be open sets such that $\Phi|_V : V \rightarrow U$ is a C^1 diffeomorphism. Let $Y_j = \Phi|_V^* X_j$ so that Y_j is a C^0 vector field on V . Write,

$$Y_j = \frac{\partial}{\partial u_j} + \sum_{k=1}^n a_j^k(u) \frac{\partial}{\partial u_k}, \quad (9.2)$$

where $a_j^k \in C(V)$. Let $A(u)$ denote the $n \times n$ matrix with j, k component $a_j^k(u)$, and let $C(u)$ denote the $n \times n$ matrix with j, k component $\sum_l u_l c_{j,l}^k \circ \Phi(u)$. We write u in polar coordinates as $u = r\theta$, $r \geq 0$.

PROPOSITION 9.1. *In the above setting, $A(u)$ satisfies the differential equation*

$$\frac{\partial}{\partial r} rA(r\theta) = -A(r\theta)^2 - C(r\theta)A(r\theta) - C(r\theta). \quad (9.3)$$

In particular, $\frac{\partial}{\partial r} rA(r\theta)$ exists in the classical sense.

LEMMA 9.2. *Proposition 9.1 holds in the special case when M is a C^∞ manifold and $X_1, \dots, X_n$ are C^∞ vector fields on M .*

Proof. When $X_1, \dots, X_n$ are C^∞ , then Φ is C^∞ and $\Phi|_V : V \rightarrow \Phi(V)$ is a C^∞ diffeomorphism. We conclude that $Y_1, \dots, Y_n$ are C^∞ vector fields. Furthermore, $[Y_j, Y_k] = \sum_l \tilde{c}_{j,k}^l Y_l$ where $\tilde{c}_{j,k}^l = c_{j,k}^l \circ \Phi$.

Note that $d\Phi(r\theta)r\frac{\partial}{\partial r} = rd\Phi(r\theta)\frac{\partial}{\partial r} = r\theta \cdot X(\Phi(r\theta))$, since $\Phi(r\theta) = e^{r(\theta \cdot X)}x$, and we are identifying X with the vector of vector fields $(X_1, \dots, X_n)$. Writing this in Cartesian coordinates, we have

$$\sum_{j=1}^n u_j \frac{\partial}{\partial u_j} = \sum_{j=1}^n u_j Y_j(u). \quad (9.4)$$

Taking the Lie bracket of (9.4) with Y_i , we obtain

$$\begin{aligned} \sum_{j=1}^n ((Y_i u_j) \partial_{u_j} + u_j [Y_i, \partial_{u_j}]) &= \sum_{j=1}^n ((Y_i u_j) Y_j + u_j [Y_i, Y_j]) \\ &= \sum_{j=1}^n \left((Y_i u_j) Y_j + u_j \sum_{l=1}^n \tilde{c}_{i,j}^l(u) Y_l \right). \end{aligned} \quad (9.5)$$

We re-write (9.5) as

$$\begin{aligned} \left(\sum_{j=1}^n u_j [\partial_{u_j}, Y_i - \partial_{u_i}] \right) + Y_i - \partial_{u_i} &= - \left(\sum_{j=1}^n ((Y_i - \partial_{u_i})(u_j)) (Y_j - \partial_{u_j}) \right) \\ &\quad - \sum_{j=1}^n \sum_{l=1}^n u_j \tilde{c}_{i,j}^l(u) Y_l. \end{aligned} \quad (9.6)$$

Plugging (9.2) into (9.6), we have

$$\begin{aligned} \sum_{j=1}^n \sum_{k=1}^n u_j (\partial_{u_j} a_i^k) \partial_{u_k} + \sum_{k=1}^n a_i^k \partial_{u_k} &= - \sum_{k=1}^n \sum_{j=1}^n a_i^j a_j^k \partial_{u_k} - \sum_{k=1}^n \sum_{j=1}^n u_j \tilde{c}_{i,j}^k \partial_{u_k} \\ &\quad - \sum_{l=1}^n \sum_{k=1}^n \sum_{j=1}^n u_j \tilde{c}_{i,j}^l a_l^k \partial_{u_k}. \end{aligned}$$

Taking the ∂_{u_k} component of the above, and writing $1 + \sum_{j=1}^n u_j \partial_{u_j} = \partial_r r$, we have

$$\partial_r r a_i^k = - \sum_{j=1}^n a_i^j a_j^k - \sum_{j=1}^n u_j \tilde{c}_{i,j}^k - \sum_{l=1}^n \left(\sum_{j=1}^n u_j \tilde{c}_{i,j}^l \right) a_l^k.$$

This is exactly (9.3) and completes the proof. $\square$

Proof of Proposition 9.1. By a classical theorem of Whitney, there is a C^∞ structure on M compatible with its C^2 structure, so we may assume M is a C^∞ manifold. Pick¹⁴ $\tilde{V} \Subset V$ and $\tilde{U} \Subset U$ open sets with $\Phi|_{\tilde{V}} : \tilde{V} \rightarrow \tilde{U}$ a C^1 diffeomorphism. Fix $u_0 \in \tilde{V}$. We will prove the result with V replaced by $B^n(u_0, \delta_0)$ for some $\delta_0 > 0$, and the result will follow as the conclusion is local.

Fix $\epsilon' \in (0, \epsilon)$ so large that $\tilde{V} \subseteq B^n(\epsilon')$. Let X_j^σ be smooth vector fields on M such that $X_j^\sigma \rightarrow X_j$ in C^1 as $\sigma \rightarrow 0$. Define

$$\Phi_\sigma(u) = e^{u_1 X_1^\sigma + \dots + u_n X_n^\sigma} x.$$

Then, for σ sufficiently small, $\Phi_\sigma(u)$ is defined for $u \in B^n(\epsilon')$, and $X_1^\sigma, \dots, X_n^\sigma$ form a basis for the tangent space at every point of a neighborhood of the closure of $\Phi_\sigma(B^n(\epsilon'))$. Thus, we may write $[X_i^\sigma, X_j^\sigma] = \sum_k c_{i,j}^{k,\sigma} X_k^\sigma$, with $c_{i,j}^{k,\sigma} \rightarrow c_{i,j}^{k,0}$ in C^0 as $\sigma \rightarrow 0$. Also, $\Phi_\sigma \rightarrow \Phi$ in $C^1(B^n(\epsilon'))$ as $\sigma \rightarrow 0$, by standard theorems.

For σ sufficiently small, $|\det d\Phi_\sigma(u_0)| \geq \frac{1}{2} |\det d\Phi(u_0)| > 0$. The Inverse Function Theorem shows that there is a $\delta_0 > 0$ (independent of σ) so that for σ small, $\Phi_\sigma|_{B^n(u_0, \delta_0)}$ is a diffeomorphism onto its image.

Define A_σ and C_σ in the obvious way on $B^n(u_0, \delta_0)$, by using the vector fields $X_1^\sigma, \dots, X_n^\sigma$. We have that $A_\sigma \rightarrow A$ and $C_\sigma \rightarrow C$ in $C^0(B^n(u_0, \delta_0))$. Furthermore, by Lemma 9.2, $\partial_r r A_\sigma = -A_\sigma^2 - C_\sigma A_\sigma - C_\sigma$. Taking the limit as $\sigma \rightarrow 0$, we find that $\partial_r r A$ exists in the classical sense and $\partial_r r A = -A^2 - CA - C$, completing the proof. $\square$

For another proof of Proposition 9.1 in the special case where ϵ is assumed to be small, see [MM13b, Appendix A].

¹⁴ Recall, $\tilde{V} \Subset V$ means that $\tilde{V}$ is a relatively compact subset of V .

9.1.2 Regularity properties. In this section, we discuss the existence, uniqueness, and regularity of solutions to (9.1) satisfying $A(0) = 0$. Some of this was done in [Str11], however we provide a complete proof here.

To facilitate the proof, we introduce a family of function spaces on $B^n(\eta)$. Throughout this section, for a matrix A , we write $|A|$ to denote the operator norm of A .

Fix $\eta > 0$, we are interested in solutions $A(x) \in C(B^n(\eta); \mathbb{M}^{n \times n})$ to (9.1) (in this section, we use the variable x in place of u). For $l \in \mathbb{N}$ set

$$\Omega_l := \{(x, h) \in B^n(\eta) \times (\mathbb{R}^n \setminus \{0\}) : x + jh \in B^n(\eta), 0 \leq j \leq l\}.$$

Note that $\Omega_0 := B^n(\eta) \times (\mathbb{R}^n \setminus \{0\})$. For $h \in \mathbb{R}^n \setminus \{0\}$ set $\Delta_h A(x) = A(x+h) - A(x)$ and $\Delta_h^l A(x) = (\Delta_h)^l A(x)$. Note that $\Delta_h^l A(x)$ is defined precisely for $(x, h) \in \Omega_l$. Without explicitly mentioning it, we will repeatedly use the fact that if $(x, h) \in \Omega_l$ and $s \in (0, 1]$, then $(sx, sh) \in \Omega_l$.

Let $\omega : (0, \infty) \rightarrow (0, \infty)$ be a non-decreasing function and for $l, m \in \mathbb{N}$ set

$$\begin{aligned} \|A\|_{C^{m,l,\omega}} &:= \sum_{|\beta| \leq m} \sum_{j=0}^l \sup_{(x,h) \in \Omega_j} \omega(|h|)^{-j} \left| \Delta_h^j \partial_x^\beta A(x) \right|, \\ C^{m,l,\omega} &:= \{A \in C^m(B^n(\eta); \mathbb{M}^{n \times n}) : \|A\|_{C^{m,l,\omega}} < \infty\}. \end{aligned}$$

Note that $C^{m,l,\omega}$ is a Banach space, and when $l = 0$, ω does not play a role.

REMARK 9.3. We are particularly interested in the following special cases

$$\begin{aligned} C^m(B^n(\eta); \mathbb{M}^{n \times n}) &= C^{m,0,\omega}, \quad C^{m,s}(B^n(\eta); \mathbb{M}^{n \times n}) = C^{m,1,\omega_s}, \\ \mathcal{C}^{m+s}(B^n(\eta); \mathbb{M}^{n \times n}) &= C^{m,2,\omega_s/2}, \end{aligned}$$

with equality of norms, where $\omega_s(h) = h^s$.

PROPOSITION 9.4. Let $C \in C(B^n(\eta); \mathbb{M}^{n \times n})$ be given with $C(0) = 0$. Suppose $|C(x)| \leq D|x|$, for $x \in B^n(\eta)$. Then, if $\eta \leq (10D)^{-1}$, there exists a unique $A \in C^0(B^n(\eta); \mathbb{M}^{n \times n})$ with $A(0) = 0$ satisfying (9.1). This unique solution satisfies:

$$|A(x)| \leq \frac{5}{8}D|x| \text{ and } |A(x)| \leq \frac{1}{16}, \quad \forall x \in B^n(\eta). \quad (9.7)$$

Furthermore, for this solution A ,

$$C \in C^{m,l,\omega} \Rightarrow A \in C^{m,l,\omega}, \quad \forall m, l, \omega,$$

and

$$\|A\|_{C^{m,l,\omega}} \leq K_{n,m,l,\omega},$$

where $K_{n,m,l,\omega}$ can be chosen to depend only on n , m , l , and an upper bound for $\|C\|_{C^{m,l,\omega}}$.

The rest of this section is devoted to the proof of Proposition 9.4. We begin with several lemmas.

LEMMA 9.5. For $j \leq m$, $k \leq l$, $C^{m,l,\omega} \hookrightarrow C^{j,k,\omega}$ and

$$\|A\|_{C^{j,k,\omega}} \leq \|A\|_{C^{m,l,\omega}}. \quad (9.8)$$

If $A, B \in C^{m,l,\omega}$, then $AB \in C^{m,l,\omega}$ and

$$\|AB\|_{C^{m,l,\omega}} \leq C_{m,l} \|A\|_{C^{m,l,\omega}} \|B\|_{C^{m,l,\omega}}, \quad (9.9)$$

where $C_{m,l}$ can be chosen to depend only on m and l .

Proof. The inclusion and inequality (9.8) follow immediately from the definitions, thus we prove only the algebra property and (9.9).

For $A, B \in C^{m,l,\omega}$ and $0 \leq j \leq l$, $|\beta| \leq m$, we have $\partial_x^\beta \Delta_h^j(AB)$ is a constant coefficient linear combination of terms of the form

$$\tau_{k_1 h} \left(\Delta_h^{j_1} \partial_x^{\beta_1} A \right) \tau_{k_2 h} \left(\Delta_h^{j_2} \partial_x^{\beta_2} B \right), \quad (9.10)$$

where $\tau_h A(x) = A(x+h)$, $j_1 + j_2 = j$, $0 \leq k_1 \leq j_2$, $0 \leq k_2 \leq j_1$, $\beta_1 + \beta_2 = \beta$. Note that, since $0 \leq k_1 \leq j_2$, $0 \leq k_2 \leq j_1$, and $j_1 + j_2 = j \leq l$, the expression in (9.10) is defined for $(x, h) \in \Omega_l$. Finally,

$$\begin{aligned} & \left| \omega(|h|)^{-j} \tau_{k_1 h} \left(\Delta_h^{j_1} \partial_x^{\beta_1} A \right) \tau_{k_2 h} \left(\Delta_h^{j_2} \partial_x^{\beta_2} B \right) \right| \\ &= \left| \tau_{k_1 h} \left(\omega(|h|)^{-j_1} \Delta_h^{j_1} \partial_x^{\beta_1} A \right) \tau_{k_2 h} \left(\omega(|h|)^{-j_2} \Delta_h^{j_2} \partial_x^{\beta_2} B \right) \right| \\ &\leq \|A\|_{C^{|\beta_1|, j_1, \omega}} \|B\|_{C^{|\beta_2|, j_2, \omega}} \leq \|A\|_{C^{m,l,\omega}} \|B\|_{C^{m,l,\omega}}, \end{aligned}$$

where the last inequality follows from (9.8). The result follows. $\square$

Define $\mathcal{T} : C(B^n(\eta); \mathbb{M}^{n \times n}) \rightarrow C(B^n(\eta); \mathbb{M}^{n \times n})$ by

$$\mathcal{T}(A)(x) = \int_0^1 -A(sx)^2 - C(sx)A(sx) - C(sx) ds.$$

The relevance of $\mathcal{T}$ is the following lemma.

LEMMA 9.6. $A \in C(B^n(\eta); \mathbb{M}^{n \times n})$ is a solution to (9.1) if and only if $\mathcal{T}(A) = A$. Also, writing $x = r\theta$, we have the following formula for $\mathcal{T}$ when $r > 0$:

$$\mathcal{T}(A)(r\theta) = \frac{1}{r} \int_0^r -A(s\theta)^2 - C(s\theta)A(s\theta) - C(s\theta) ds. \quad (9.11)$$

Proof. (9.11) follows from a straightforward change of variables in the definition of $\mathcal{T}$. That $A \in C(B^n(\eta); \mathbb{M}^{n \times n})$ is a solution to (9.1) if and only if $\mathcal{T}(A) = A$ follows from (9.11). $\square$

LEMMA 9.7. If $C \in C^{m,l,\omega}$, then $\mathcal{T} : C^{m,l,\omega} \rightarrow C^{m,l,\omega}$.

Proof. Let $A \in C^{m,l,\omega}$. We wish to show $\mathcal{T}(A) \in C^{m,l,\omega}$. Set $B := -A^2 - CA - C$. By Lemma 9.5, $B \in C^{m,l,\omega}$. We wish to show $\int_0^1 B(sx) ds \in C^{m,l,\omega}$.

Let $0 \leq j \leq l$, $|\beta| \leq m$. Consider,

$$\begin{aligned} \left| \Delta_h^j \partial_x^\beta \int_0^1 B(sx) ds \right| &= \left| \int_0^1 s^{|\beta|} (\Delta_{sh}^j \partial_x^\beta B)(sx) ds \right| \leq \int_0^1 s^{|\beta|} \omega(s|h|)^j \|B\|_{C^{m,l,\omega}} ds \\ &\leq \omega(|h|)^j \|B\|_{C^{m,l,\omega}} (|\beta| + 1)^{-1}, \end{aligned}$$

where we have used that ω is non-decreasing. The result follows. $\square$

LEMMA 9.8. (Izzo's contraction mapping principle [Izz99]). Suppose (X, d) is a metric space and $\{\mathcal{Q}_a\}_{a=0}^\infty$ is a sequence of contractions on X for which there exists $c < 1$ with

$$d(\mathcal{Q}_a(x), \mathcal{Q}_a(y)) \leq cd(x, y), \quad \forall x, y \in X, a \in \mathbb{N}.$$

Suppose $\exists x_\infty \in X$ with $\lim_{a \rightarrow \infty} \mathcal{Q}_a(x_\infty) = x_\infty$. Let $x_0 \in X$ be arbitrary, and define x_a recursively by $x_{a+1} = \mathcal{Q}_a(x_a)$. Then $\lim_{a \rightarrow \infty} x_a = x_\infty$.

Proof. We include a slightly modified version of the proof in [Izz99]. For each $a \in \mathbb{N}$,

$$\begin{aligned} d(x_{a+1}, x_\infty) &= d(\mathcal{Q}_a(x_a), x_\infty) \leq d(\mathcal{Q}_a(x_a), \mathcal{Q}_a(x_\infty)) + d(\mathcal{Q}_a(x_\infty), x_\infty) \\ &\leq cd(x_a, x_\infty) + d(\mathcal{Q}_a(x_\infty), x_\infty). \end{aligned} \quad (9.12)$$

First we claim that the sequence $d(x_a, x_\infty)$ is bounded. Since $\mathcal{Q}_a(x_\infty) \rightarrow x_\infty$, $\exists N$, $a \geq N \Rightarrow d(\mathcal{Q}_a(x_\infty), x_\infty) < 1 - c$. Suppose $d(x_a, x_\infty)$ is not bounded; then $\exists a \geq N$ with $\max\{d(x_a, x_\infty), 1\} \leq d(x_{a+1}, x_\infty)$. Applying this to (9.12), we have $d(x_{a+1}, x_\infty) \leq cd(x_a, x_\infty) + d(\mathcal{Q}_a(x_\infty), x_\infty) < cd(x_{a+1}, x_\infty) + 1 - c$. And so $d(x_{a+1}, x_\infty) < 1 \leq d(x_{a+1}, x_\infty)$, a contradiction. Thus the sequence $d(x_a, x_\infty)$ is bounded.

Since $\mathcal{Q}_a(x_\infty) \rightarrow x_\infty$, (9.12) implies

$$\limsup_{a \rightarrow \infty} d(x_a, x_\infty) \leq c \limsup_{a \rightarrow \infty} d(x_a, x_\infty).$$

Since $\limsup_{a \rightarrow \infty} d(x_a, x_\infty) < \infty$, this gives $\limsup_{a \rightarrow \infty} d(x_a, x_\infty) = 0$, completing the proof. $\square$

We now turn to Proposition 9.4. We begin with uniqueness. Suppose $A_1, A_2 \in C(B^n(\eta); \mathbb{M}^{n \times n})$ are two solutions to (9.1) with $A_1(0) = A_2(0) = 0$. By Lemma 9.6 we have $\mathcal{T}(A_1) = A_1$, $\mathcal{T}(A_2) = A_2$. We first claim that $|A_j(x)| = O(|x|)$ for $j = 1, 2$; we prove this for A_1 and the same is true for A_2 by symmetry. Set $F(r) = \sup_{|x| \leq r} |A_1(x)|$, note that $F : [0, \eta) \rightarrow \mathbb{R}$ is continuous, increasing, and $F(0) = 0$. Since $\mathcal{T}(A_1) = A_1$ and $|C(sx)| \leq Ds|x|$ by assumption, we have

$$|A_1(x)| \leq \int_0^1 F(s|x|)^2 + Ds|x|F(s|x|) + Ds|x| ds \leq F(|x|)^2 + \frac{1}{2}D|x|F(|x|) + \frac{1}{2}D|x|.$$

And so $F(r) \leq F(r)^2 + \frac{1}{2}DrF(r) + \frac{1}{2}Dr$, and thus $F(r)(1 - F(r)) \leq \frac{1}{2}DrF(r) + \frac{1}{2}Dr$. Taking r so small that $F(r) \leq \frac{1}{2}$, we have for such r , $F(r) \leq \frac{3}{2}Dr$. Thus $|A_1(x)| = O(|x|)$.

Writing x in polar coordinates $x = r\theta$ and using (9.11) we have for $r > 0$,

$$\begin{aligned} & |r(A_1(r\theta) - A_2(r\theta))| \\ & \leq \int_0^r |s(A_1(s\theta) - A_2(s\theta))| (s^{-1}|A_1(s\theta)| + s^{-1}|A_2(s\theta)| + s^{-1}|C(s\theta)|) ds. \end{aligned}$$

Using that $|A_1(s\theta)|, |A_2(s\theta)|, |C(s\theta)| = O(s)$, the integral form of Grönwall's inequality shows that $A_1(r\theta) = A_2(r\theta)$ for $r > 0$ and therefore $A_1 = A_2$. This completes the proof of uniqueness.

We now turn to existence for which we use the contraction mapping principle. Let

$$\mathcal{M} := \left\{ A \in C^0(B^n(\eta); \mathbb{M}^{n \times n}) \mid A(0) = 0, \sup_{0 \neq x \in B^n(\eta)} \frac{1}{|x|} |A(x)| < \infty, \right. \\ \left. \sup_{x \in B^n(\eta)} |A(x)| \leq \frac{1}{10} \right\}.$$

We give $\mathcal{M}$ the metric

$$d(A, B) := \sup_{0 \neq x \in B^n(\eta)} \frac{1}{|x|} |A(x) - B(x)|.$$

With this metric, $\mathcal{M}$ is a complete metric space.

LEMMA 9.9. $\mathcal{T} : \mathcal{M} \rightarrow \mathcal{M}$ and $\forall A, B \in \mathcal{M}$, $d(\mathcal{T}(A), \mathcal{T}(B)) \leq \frac{1}{5}d(A, B)$. Also, $d(\mathcal{T}(0), 0) \leq D/2$.

Proof. Let $A \in \mathcal{M}$. For $x \in B^n(\eta)$,

$$\begin{aligned} |\mathcal{T}(A)(x)| & \leq \int_0^1 \|A\|_{C^0}^2 + Ds|x|\|A\|_{C^0} + Ds|x| ds \leq \frac{1}{100} + \frac{D}{2}\eta\frac{1}{10} + \frac{D}{2}\eta \\ & \leq \frac{1}{100} + \frac{1}{200} + \frac{1}{20} \leq \frac{1}{10}. \end{aligned} \tag{9.13}$$

Also,

$$\frac{1}{|x|} |\mathcal{T}(0)(x)| \leq \frac{1}{|x|} \int_0^1 Ds|x| ds \leq \frac{1}{2}D, \tag{9.14}$$

and so $\mathcal{T}(0) \in \mathcal{M}$ with $d(\mathcal{T}(0), 0) \leq D/2$.

Finally, for $A, B \in \mathcal{M}$, $0 \neq x \in B^n(\eta)$,

$$\begin{aligned} \frac{1}{|x|} |\mathcal{T}(A)(x) - \mathcal{T}(B)(x)| &\leq \frac{1}{|x|} \int_0^1 |A(sx) - B(sx)| (|A(sx)| + |B(sx)| + |C(sx)|) ds \\ &\leq \frac{1}{|x|} \int_0^1 s|x| d(A, B) \left(\frac{1}{5} + Ds|x| \right) ds \leq \int_0^1 sd(A, B) \left(\frac{1}{5} + \frac{s}{10} \right) ds \\ &\leq d(A, B) \left(\frac{1}{10} + \frac{1}{30} \right) \leq \frac{1}{5} d(A, B). \end{aligned} \quad (9.15)$$

Putting $0 = B$ in (9.15) and using (9.14) shows $\sup_{0 \neq x \in B^n(\eta)} \frac{1}{|x|} |\mathcal{T}(A)(x)| < \infty$. Combining this with (9.13) shows $\mathcal{T} : \mathcal{M} \rightarrow \mathcal{M}$. Further, (9.15) with arbitrary $A, B \in \mathcal{M}$ shows $d(\mathcal{T}(A), \mathcal{T}(B)) \leq \frac{1}{5} d(A, B)$, and this completes the proof. $\square$

By Lemma 9.9, $\mathcal{T} : \mathcal{M} \rightarrow \mathcal{M}$ is a strict contraction, and the contraction mapping principle applies to show that if $A_0 = 0$, $A_a = \mathcal{T}(A_{a-1})$, $a \geq 1$, then $A_a \rightarrow A_\infty$ in $\mathcal{M}$, where $\mathcal{T}(A_\infty) = A_\infty$. A_∞ is the desired solution to (9.1).

Also, for $a \in \mathbb{N} \cup \{\infty\}$ we have, using Lemma 9.9,

$$\frac{1}{|x|} |A_a(x)| \leq d(A_a, 0) \leq \sum_{b=0}^{a-1} d(\mathcal{T}^{b+1}(0), \mathcal{T}^b(0)) \leq \sum_{b=0}^{a-1} 5^{-b} d(\mathcal{T}(0), 0) \leq \frac{5}{8} D. \quad (9.16)$$

In particular, for $x \in B^n(\eta)$, $|A_\infty(x)| \leq \frac{5}{8} D|x|$. Also, since $\eta \leq (10D)^{-1}$, it follows that $|A_\infty(x)| \leq \frac{1}{16}$; this establishes (9.7).

It remains to prove the regularity properties of A_∞ , in terms of the regularity of C . For the remainder of this section, $K_{n,m,l,\omega}$ is a constant which can be chosen to depend only on n, m, l , and an upper bound for $\|C\|_{C^{m,l,\omega}}$. This constant may change from line to line.

To complete the proof of Proposition 9.4, we will prove the following when $C \in C^{m,l,\omega}$:

- $A_a \rightarrow A_\infty$ in $C^{m,l,\omega}$.
- $\|A_\infty\|_{C^{m,l,\omega}} \leq K_{n,m,l,\omega}$.

We prove the above two properties by induction on m, l . The base case, $m = l = 0$, was just proved above (since $C^{0,0,\omega} = C^0(B^n(\eta); \mathbb{M}^{n \times n})$).

Fix (m, l) . We assume we have the above for all (k, j) with $0 \leq k \leq m$, $0 \leq j \leq l$, and $(k, j) \neq (m, l)$, and we assume $C \in C^{m,l,\omega}$. Since for $0 \leq k \leq m$, $0 \leq j \leq l$, $C^{m,l,\omega} \hookrightarrow C^{k,j,\omega}$ (Lemma 9.5), the inductive hypothesis shows for such (k, j) with $(k, j) \neq (m, l)$, $A_a \rightarrow A_\infty$ in $C^{k,j,\omega}$ and $\|A_\infty\|_{C^{k,j,\omega}} \leq K_{n,k,j,\omega}$.

We define a Banach space $X_{\omega,l}$ as follows:

- $X_{\omega,0} = C(B^n(\eta); \mathbb{M}^{n \times n})$, with the usual norm.
- For $l > 0$, $X_{\omega,l} = \{B(x, h) \in C(\Omega_l; \mathbb{M}^{n \times n}) : \|B\|_{X_{\omega,l}} < \infty\}$, where $\|B\|_{X_{\omega,l}} := \sup_{(x,h) \in \Omega_l} \omega(|h|)^{-l} |B(x, h)|$.

Fix $|\beta| = m$. We will show (under our inductive hypothesis)

- (i) $A_a \in C^{m,l,\omega}$, $\forall a \in \mathbb{N}$.
- (ii) $\Delta_h^l \partial_x^\beta A_a(x) \in X_{\omega,l}$, $\forall a \in \mathbb{N}$.
- (iii) $\exists B_\infty \in X_{\omega,l}$ such that $\Delta_h^l \partial_x^\beta A_a \xrightarrow{a \rightarrow \infty} B_\infty$ in $X_{\omega,l}$.
- (iv) $\|B_\infty\|_{X_{\omega,l}} \leq K_{n,m,l,\omega}$.

First we see why the above completes the proof. We already know from our inductive hypothesis that

$$\sup_{(x,h) \in \Omega_j} \omega(|h|)^{-j} \left| \Delta_h^j \partial_x^\alpha (A_a - A_\infty) \right| \xrightarrow{a \rightarrow \infty} 0, \quad (9.17)$$

for $0 \leq j \leq l$, $|\alpha| \leq m$ with $(j, |\alpha|) \neq (l, m)$, and that $\|A_\infty\|_{C^{j,k,\omega}} \leq K_{n,k,j,\omega}$ for $0 \leq k \leq m$, $0 \leq j \leq l$, $(j, k) \neq (l, m)$. Thus, that $A_a \rightarrow A_\infty$ in $C^{m,l,\omega}$ will follow from (9.17) for $(j, |\alpha|) = (l, m)$ and the fact that $A_a \in C^{m,l,\omega}$. If $l = 0$, (i) implies $A_a \in C^m$ and (iii) implies $\partial_x^\beta A_a \rightarrow B_\infty$ in the supremum norm. Since $A_a \rightarrow A_\infty$ in C^0 , we have $\partial_x^\beta A_\infty = B_\infty$. (iv) implies the desired bound on $\partial_x^\beta A_\infty$. Since β is arbitrary with $|\beta| = m$, we conclude $A_\infty \in C^m$, with $\|A_\infty\|_{C^{m,0,\omega}} \leq K_{n,m,0,\omega}$, and $A_a \rightarrow A_\infty$ in $C^{m,0,\omega}$, as desired.

If $l \geq 1$, then we already know $A_a \rightarrow A_\infty$ in $C^m(B^n(\eta); \mathbb{M}^{n \times n})$, by the inductive hypothesis. Thus

$$\Delta_h^l \partial_x^\beta A_a(x) \rightarrow \Delta_h^l \partial_x^\beta A_\infty, \quad \text{pointwise.}$$

Hence, $\Delta_h^l \partial_x^\beta A_\infty(x) = B_\infty(x, h)$. Since β was arbitrary with $|\beta| = m$, (iii) shows $A_a \rightarrow A_\infty$ in $C^{m,l,\omega}$ and (iv) shows $\|A_\infty\|_{C^{m,l,\omega}} \leq K_{n,m,l,\omega}$.

Having shown them to be sufficient, we turn to proving (i) to (iv). Recall, we have fixed β with $|\beta| = m$. Since $A_a = \mathcal{T}^a(0)$, (i) follows from Lemma 9.7. (ii) is an immediate consequence of (i). Thus, it remains only to prove (iii) and (iv). We will do this by applying Lemma 9.8. To begin, we need a few preliminary lemmas.

LEMMA 9.10. *Fix $m_1, l_1, m_2, l_2, j_1, j_2 \in \mathbb{N}$ and set $l = l_1 + l_2$ and suppose $j_1 + l_1, j_2 + l_2 \leq l$. Let β_1 and β_2 be multi-indices with $|\beta_1| = m_1$ and $|\beta_2| = m_2$. Then, the bilinear map for $A_1 \in C^{m_1,l_1,\omega}$, $A_2 \in C^{m_2,l_2,\omega}$ given by*

$$(A_1, A_2) \mapsto \left(\tau_{j_1 h} \Delta_h^{l_1} \partial_x^{\beta_1} A_1 \right)(x) \left(\tau_{j_2 h} \Delta_h^{l_2} \partial_x^{\beta_2} A_2 \right)(x). \quad (9.18)$$

is a continuous map $C^{m_1,l_1,\omega} \times C^{m_2,l_2,\omega} \rightarrow X_{\omega,l}$, and the norm of this map is ≤ 1 . Here, $\tau_h A(x) = A(x + h)$.

Proof. The restriction $j_1 + l_1, j_2 + l_2 \leq l$, ensures that the expression in (9.18) is defined for $(x, h) \in \Omega_l$. With this in mind, the result follows immediately from the definitions. $\square$

For an element $B \in X_{\omega,l}$ we often write $B(x, h)$. When $l \geq 1$, the meaning of this is obvious. For $l = 0$ this is to be interpreted as $B(x)$.

LEMMA 9.11. For $B(x, h) \in X_{\omega, l}$ and $d \geq 1$ the map

$$B \mapsto \int_0^1 s^d B(sx, sh) ds$$

is continous $X_{\omega, l} \rightarrow X_{\omega, l}$ and has norm ≤ 1 .

Proof. This is clear from the definitions. $\square$

For $A_1, A_2 \in C^{m, \omega, l}$, we have

$$\Delta_h^l \partial_x^\beta (A_1 A_2)(x) = (\Delta_h^l \partial_x^\beta A_1)(x) A_2(x + lh) + A_1(x) (\Delta_h^l \partial_x^\beta A_2)(x) + R_{\beta, l}(A_1, A_2)(x),$$

where $R_{\beta, l}(A_1, A_2)(x, h)$ is a constant coefficient linear combination (depending only on β and l) of terms of the form

$$\left(\tau_{j_1 h} \Delta_h^{l_1} \partial_x^{\beta_1} A_1 \right) (x) \left(\tau_{j_2 h} \Delta_h^{l_2} \partial_x^{\beta_2} A_2 \right) (x),$$

where $0 \leq j_1 \leq l_2$, $0 \leq j_2 \leq l_1$, $l_1 + l_2 = l$, $\beta_1 + \beta_2 = \beta$, and $l_1 + |\beta_1|, l_2 + |\beta_2| > 0$.

LEMMA 9.12. We have the following limits in $X_{\omega, l}$:

$$\int_0^1 s^{|\beta|} R_{\beta, l}(A_a, A_a)(sx, sh) ds \xrightarrow{a \rightarrow \infty} \int_0^1 s^{|\beta|} R_{\beta, l}(A_\infty, A_\infty)(sx, sh) ds. \quad (9.19)$$

$$\int_0^1 s^{|\beta|} R_{\beta, l}(C, A_a)(sx, sh) ds \xrightarrow{a \rightarrow \infty} \int_0^1 s^{|\beta|} R_{\beta, l}(C, A_\infty)(sx, sh) ds. \quad (9.20)$$

$$\int_0^1 s^{|\beta|} \left(\Delta_{sh}^l \partial_x^\beta C \right) (sx) A_a(s(x + lh)) ds \xrightarrow{a \rightarrow \infty} \int_0^1 s^{|\beta|} \left(\Delta_{sh}^l \partial_x^\beta C \right) (sx) A_\infty(s(x + lh)) ds. \quad (9.21)$$

And for any $B(x, h) \in X_{\omega, l}$,

$$\int_0^1 s^{|\beta|} B(sx, sh) A_a(s(x + lh)) ds \xrightarrow{a \rightarrow \infty} \int_0^1 s^{|\beta|} B(s, x) A_\infty(s(x + lh)) ds. \quad (9.22)$$

$$\int_0^1 s^{|\beta|} A_a(sx) B(sx, sh) ds \xrightarrow{a \rightarrow \infty} \int_0^1 s^{|\beta|} A_\infty(sx) B(sx, sh) ds. \quad (9.23)$$

Proof. Recall, we are assuming $C \in C^{m, \omega, l}$ and our inductive hypothesis implies $A_a \rightarrow A_\infty$ in $C^{k, j, \omega}$ with $0 \leq k \leq m$, $0 \leq j \leq l$, and $(k, j) \neq (m, l)$. Using this and Lemmas 9.10 and 9.11, (9.19), (9.20), and (9.21) follow immediately. (9.22) and (9.23) follow from the fact that $A_a \rightarrow A_\infty$ in $C^0(B^n(\eta))$ and a straightforward estimate. $\square$

LEMMA 9.13.

$$\begin{aligned} \left\| \int_0^1 s^{|\beta|} R_{\beta,l}(A_\infty, A_\infty)(sx, sh) ds \right\|_{X_{\omega,l}} &\leq K_{n,m,l,\omega}. \\ \left\| \int_0^1 s^{|\beta|} R_{\beta,l}(C, A_\infty)(sx, sh) ds \right\|_{X_{\omega,l}} &\leq K_{n,m,l,\omega}. \\ \left\| \int_0^1 s^{|\beta|} \left(\Delta_{sh}^l \partial_x^\beta C \right) (sx) A_\infty(s(x+lh)) ds \right\|_{X_{\omega,l}} &\leq K_{n,m,l,\omega}. \end{aligned}$$

Proof. This follows from the inductive hypothesis and Lemmas 9.10 and 9.11. $\square$

For $a \in \mathbb{N} \cup \{\infty\}$, $B \in X_{\omega,l}$, define

$$\begin{aligned} \mathcal{Q}_a(B)(x, h) = \int_0^1 -s^{|\beta|} \Big[&B(sx, sh)A_a(s(x+lh)) + A_a(sx)B(sx, sh) \\ &+ C(sx)B(sx, sh) + \left(\Delta_{sh}^l \partial_x^\beta C \right) (sx)A_a(s(x+lh)) + \left(\Delta_{sh}^l \partial_x^\beta C \right) (sx) \\ &+ R_{\beta,l}(A_a, A_a)(sx, sh) + R_{\beta,l}(C, A_a)(sx, sh) \Big] ds \end{aligned}$$

LEMMA 9.14. For $a \in \mathbb{N} \cup \{\infty\}$, $\mathcal{Q}_a : X_{\omega,l} \rightarrow X_{\omega,l}$ and satisfies

$$\|\mathcal{Q}_a(B) - \mathcal{Q}_a(B')\|_{X_{\omega,l}} \leq \frac{1}{8} \|B - B'\|_{X_{\omega,l}}. \quad (9.24)$$

Furthermore, $\forall B \in X_{\omega,l}$, $\lim_{a \rightarrow \infty} \mathcal{Q}_a(B) = \mathcal{Q}_\infty(B)$. Finally, $\|\mathcal{Q}_\infty(0)\|_{X_{\omega,l}} \leq K_{n,m,l,\omega}$.

Proof. That $\mathcal{Q}_a : X_{\omega,l} \rightarrow X_{\omega,l}$ follows from Lemmas 9.10 and 9.11, the inductive hypothesis, and the fact that $A_a \in C^0(B^n(\eta))$, $\forall a \in \mathbb{N} \cup \{\infty\}$.

That $\lim_{a \rightarrow \infty} \mathcal{Q}_a(B) = \mathcal{Q}_\infty(B)$ follows from Lemma 9.12 and $\|\mathcal{Q}_\infty(0)\|_{X_{\omega,l}} \leq K_{n,m,l,\omega}$ follows from Lemmas 9.11 and 9.13.

Thus we need only show (9.24). We have, using (9.16), for $(x, h) \in \Omega_l$, $a \in \mathbb{N} \cup \{\infty\}$,

$$\begin{aligned} &|\mathcal{Q}_a(B)(x, h) - \mathcal{Q}_a(B')(x, h)| \\ &\leq \int_0^1 s^{|\beta|} (|A_a(s(x+lh))| + |A_a(sx)| + |C(sx)|) |B(sx, sh) - B'(sx, sh)| ds \\ &\leq \int_0^1 \left(\frac{5}{8} Ds|x+lh| + \frac{5}{8} Ds|x| + Ds|x| \right) \omega(s|h|)^l \|B - B'\|_{X_{\omega,l}} ds \\ &\leq \|B - B'\|_{X_{\omega,l}} \int_0^1 \frac{9}{4} Ds\eta\omega(|h|)^l ds \\ &\leq \frac{1}{8} \omega(|h|)^l \|B - B'\|_{X_{\omega,l}}, \end{aligned}$$

completing the proof of (9.24), and therefore the proof of the lemma. $\square$

For $a \in \mathbb{N}$, define $B_a(x, h) := \Delta_h^l \partial_x^\beta A_a(x)$; note that $B_a \in X_{\omega, l}$ since $A_a \in C^{m, \omega, l}$. Also, $B_{a+1}(x, h) = \Delta_h^l \partial_x^\beta \mathcal{T}(A_a)(x) = \mathcal{Q}_a(B_a)(x, h)$.

Since $\mathcal{Q}_\infty$ is a strict contraction (Lemma 9.14), there exists a unique fixed point $B_\infty \in X_{\omega, l}$. Since $\mathcal{Q}_a(B_\infty) \rightarrow \mathcal{Q}_\infty(B_\infty) = B_\infty$, by Lemma 9.14, Lemma 9.8 shows $B_a \rightarrow B_\infty$ in $X_{\omega, l}$. Since $B_a(x, h) = \Delta_h^l \partial_x^\beta A_a(x)$, this proves (iii).

Finally, to prove (iv) note that B_∞ is the fixed point of the strict contraction $\mathcal{Q}_\infty$. Thus, $\mathcal{Q}_\infty^a(0) \rightarrow B_\infty$. Hence,

$$\begin{aligned} \|B_\infty\|_{X_{\omega, l}} &\leq \sum_{a=0}^{\infty} \|\mathcal{Q}_\infty^{a+1}(0) - \mathcal{Q}_\infty^a(0)\|_{X_{\omega, l}} \leq \sum_{a=0}^{\infty} 8^{-a} \|\mathcal{Q}_\infty(0) - 0\|_{X_{\omega, l}} = \frac{8}{7} \|\mathcal{Q}_\infty(0)\|_{X_{\omega, l}} \\ &\leq K_{n, m, l, \omega}, \end{aligned}$$

where the last inequality follows from Lemma 9.14. This completes the proof.

9.2 An inverse function theorem. We require a quantitative version of a special case of the Inverse Function Theorem that does not follow from the standard statement of the theorem, though we will be able to achieve it by keeping track of some constants in a standard proof. We present it here.

Fix $\eta > 0$ and let $Y_1, \dots, Y_n \in C^1(B^n(\eta); \mathbb{R}^n)$ be vector fields on $B^n(\eta)$ and suppose they satisfy

$$\inf_{u \in B^n(\eta)} |\det(Y_1(u) \cdots Y_n(u))| \geq c_0 > 0.$$

Take $C_0 > 0$ so that $\|Y_j\|_{C^1(B^n(\eta); \mathbb{R}^n)} \leq C_0$, $\forall j$. Define

$$\Psi_u(v) := e^{v_1 Y_1 + \cdots + v_n Y_n} u.$$

PROPOSITION 9.15. *There exist $\kappa = \kappa(C_0, c_0, n) > 0$ and $\Delta_0 = \Delta_0(C_0, c_0, n, \eta) > 0$ such that $\forall \delta \in (0, \Delta_0]$, $u \in B^n(\kappa\delta)$, $v \mapsto \Psi_u(v)$ is defined and injective on $v \in B^n(\delta)$. Furthermore, $B^n(\kappa\delta) \subseteq \Psi_u(B^n(\delta))$.*

The rest of this section is devoted to the proof of Proposition 9.15; for a closely related result see [MM13b, Theorem 4.5].

LEMMA 9.16. *Let $\delta_0 > 0$, $F \in C^1(B^n(\delta_0); \mathbb{R}^n)$, and suppose $dF(0)$ is nonsingular and $\sup_{x \in B^n(\delta_0)} \|dF(0)^{-1} dF(x) - I\|_{\mathbb{M}^{n \times n}} \leq \frac{1}{2}$. Then $F(B^n(\delta_0)) \subseteq \mathbb{R}^n$ is open and $F : B^n(\delta_0) \rightarrow F(B^n(\delta_0))$ is a C^1 diffeomorphism. Furthermore, $F(B^n(\delta_0)) \supseteq B^n(F(0), \kappa\delta_0)$ where*

$$\kappa := \|d(F^{-1})\|_{C^0(F(B^n(\delta_0)); \mathbb{M}^{n \times n})}^{-1} \geq c_n |\det dF(0)| \|F\|_{C^1(B^n(\delta_0); \mathbb{R}^n)}^{-(n-1)}, \quad (9.25)$$

and $c_n > 0$ can be chosen to depend only on n .

Proof. We first show F is injective. Fix $y \in \mathbb{R}^n$ and set $\phi(x) = x + dF(0)^{-1}(y - F(x))$. Note that $F(x) = y \Leftrightarrow \phi(x) = x$. Also, $\forall x \in B^n(\delta_0)$, $\|d\phi(x)\|_{\mathbb{M}^{n \times n}} \leq \|I - dF(0)^{-1}dF(x)\|_{\mathbb{M}^{n \times n}} \leq \frac{1}{2}$. Hence $|\phi(x_1) - \phi(x_2)| \leq \frac{1}{2}|x_1 - x_2|$. Hence, there is at most one solution of $\phi(x) = x$, and therefore at most one solution of $F(x) = y$, proving that F is injective.

Since $\|dF(0)^{-1}dF(x) - I\|_{\mathbb{M}^{n \times n}} \leq \frac{1}{2}$, $\forall x \in B^n(\delta_0)$, it follows that $dF(x)$ is invertible $\forall x \in B^n(\delta_0)$. Combining this with the fact that F is injective, the Inverse Function Theorem shows $F(B^n(\delta_0))$ is open and $F : B^n(\delta_0) \rightarrow F(B^n(\delta_0))$ is a C^1 diffeomorphism.

Next we prove the bound for κ given in (9.25). In what follows, we use $A \lesssim B$ to denote $A \leq C_n B$, where C_n can be chosen to depend only on n . Since $\|dF(0)^{-1}dF(x) - I\| \leq \frac{1}{2}$, by assumption,

$$\inf_{x \in B^n(\eta)} |\det dF(x)| \gtrsim |\det dF(0)|. \quad (9.26)$$

Also, $\forall x \in B^n(\delta_0)$,

$$\|(dF(x))^{-1}\|_{\mathbb{M}^{n \times n}} \lesssim |\det dF(x)|^{-1} \|dF\|_{C^0(B^n(\delta_0); \mathbb{M}^{n \times n})}^{n-1},$$

as can be seen via the cofactor representation $dF(x)^{-1}$. Hence,

$$\sup_{x \in B^n(\delta_0)} \|(dF(x))^{-1}\|_{\mathbb{M}^{n \times n}} \lesssim \left(\inf_{y \in B^n(\delta_0)} |\det dF(y)| \right)^{-1} \|dF\|_{C^0(B^n(\delta_0); \mathbb{M}^{n \times n})}^{n-1},$$

and therefore

$$\begin{aligned} \|d(F^{-1})\|_{C^0(F(B^n(\delta_0)); \mathbb{M}^{n \times n})} &\lesssim \left(\inf_{x \in B^n(\delta_0)} |\det dF(x)| \right)^{-1} \|dF\|_{C^0(B^n(\delta_0); \mathbb{M}^{n \times n})}^{n-1} \\ &\lesssim \left(\inf_{x \in B^n(\delta_0)} |\det dF(x)| \right)^{-1} \|F\|_{C^1(B^n(\delta_0); \mathbb{R}^n)}^{n-1}. \end{aligned} \quad (9.27)$$

Combining (9.26) and (9.27) yields (9.25).

Finally, we prove $F(B^n(\delta_0)) \supseteq B(F(0), \kappa\delta_0)$. Take $\epsilon > 0$ to be the largest ϵ so that $B^n(F(0), \epsilon) \subseteq F(B^n(\delta_0))$ (note that $\epsilon > 0$ by the Inverse Function Theorem). The proof will be complete once we show $\epsilon \geq \delta_0\kappa$. Suppose, for contradiction, $\epsilon < \delta_0\kappa$. We have, by the Mean Value Theorem,

$$F^{-1}(B(F(0), \epsilon)) \subseteq B(0, \epsilon \|dF^{-1}\|_{C^0(F(B^n(\delta_0)); \mathbb{M}^{n \times n})}).$$

Thus, if $\epsilon < \kappa\delta_0$, $F^{-1}(B(F(0), \epsilon)) \Subset B(0, \delta_0)$, which contradicts the choice of ϵ and completes the proof. $\square$

LEMMA 9.17. *Let Y_j , C_0 , n , η , and Ψ be as in Proposition 9.15. There exists $\delta_1 = \delta_1(C_0, n, \eta) > 0$ such that $\forall u \in B^n(\eta/2)$, Ψ_u is defined on $B^n(\delta_1)$ and satisfies*

$$\|\Psi_u\|_{C^1(B^n(\delta_1); \mathbb{R}^n)} \leq C(C_0, n) \quad (9.28)$$

and $\forall u \in B^n(\eta/2), v \in B^n(\delta_1)$,

$$\|d_v \Psi_u(v) - d_v \Psi_u(0)\|_{\mathbb{M}^{n \times n}} \leq C(C_0, n)|v|, \quad (9.29)$$

where $C(C_0, n)$ can be chosen to depend only on C_0 and n .

Proof. The existence of $\delta_1 > 0$ so that $\forall u \in B^n(\eta/2)$, $\Psi_u(v)$ is defined and (9.28) holds are classical theorems from ODEs. Thus, we prove only (9.29). We write $A \lesssim B$ for $A \leq CB$ where C can be chosen to depend only on C_0 and n . We use the equation $\partial_r \Psi_u(rv) = v \cdot Y(\Psi_u(rv))$, and so

$$\Psi_u(v) = \int_0^1 v \cdot Y(\Psi_u(sv)) ds.$$

Since $d_v \Psi_u(0) = (Y_1(u)| \cdots | Y_n(u))$, we have $\forall u \in B^n(\eta/2), v \in B^n(\delta_1)$

$$\Psi_u(v) - (d_v \Psi_u(0))v = \int_0^1 v \cdot (Y(\Psi_u(sv)) - Y(\Psi_u(0))) ds.$$

Applying d_v to the above equation and using the chain rule, we have $\forall u \in B^n(\eta/2), v \in B^n(\delta_1)$,

$$\begin{aligned} & \|d_v \Psi_u(v) - d_v \Psi_u(0)\|_{\mathbb{M}^{n \times n}} \\ &= \left\| \int_0^1 (Y(\Psi_u(sv)) - Y(\Psi_u(0))) + sv^\top dY(\Psi_u(sv))(d_v \Psi_u)(sv) ds \right\|_{\mathbb{M}^{n \times n}} \\ &\lesssim |v| \|Y \circ \Psi_u\|_{C^1(B^n(\delta_0); \mathbb{M}^{n \times n})} + |v| \|Y\|_{C^1(B^n(\eta); \mathbb{M}^{n \times n})} \|\Psi_u\|_{C^1(B^n(\delta_1); \mathbb{R}^n)} \\ &\lesssim |v| \|Y\|_{C^1(B^n(\eta); \mathbb{M}^{n \times n})} \|\Psi_u\|_{C^1(B^n(\delta_1); \mathbb{R}^n)} \lesssim |v|, \end{aligned}$$

where we have written $Y(u)$ for the matrix valued function $(Y_1(u)| \cdots | Y_n(u))$ and used (9.28). This completes the proof. $\square$

Proof of Proposition 9.15. In what follows we write $A \lesssim B$ for $A \leq CB$, where C can be chosen to depend only on n , C_0 , and c_0 , and write $A \lesssim_\eta B$ if C can also depend on η . By taking $\delta_1 \gtrsim_\eta 1$ as in Lemma 9.17, for all $u \in B^n(\eta/2)$, $v \in B^n(\delta_1)$, $\Psi_u(v)$ is defined. For such u , since $|\det d\Psi_u(0)| = |\det(Y_1(u)| \cdots | Y_n(u))| \gtrsim 1$ and using (9.28), we have $\|d_v \Psi_u(0)^{-1}\|_{\mathbb{M}^{n \times n}} \lesssim 1$. Hence, using (9.29), for $u \in B^n(\eta/2)$, $v \in B^n(\delta_1)$,

$$\|d_v \Psi_u(0)^{-1} d_v \Psi_u(v) - I\|_{\mathbb{M}^{n \times n}} \lesssim \|d_v \Psi_u(v) - d_v \Psi_u(0)\|_{\mathbb{M}^{n \times n}} \lesssim |v|.$$

Thus, if $\delta_2 \gtrsim_\eta 1$ is sufficiently small, for all $u \in B^n(\eta/2)$, $v \in B^n(\delta_2)$,

$$\|d_v \Psi_u(0)^{-1} d_v \Psi_u(v) - I\|_{\mathbb{M}^{n \times n}} \leq \frac{1}{2}.$$

By Lemma 9.16, for $|u| \leq \eta/2$, $\Psi_u : B^n(\delta_2) \rightarrow \Psi_u(B^n(\delta_2))$ is a C^1 diffeomorphism, and if we set $\kappa := \frac{1}{2} \inf_{|u| < \eta/2} \|d(\Psi_u^{-1})\|_{C^0(\Psi_u(B^n(\delta_2)); \mathbb{M}^{n \times n})}^{-1}$, we have $\kappa \gtrsim 1$ (also by

Lemma 9.16). Notice the extra factor of $1/2$ in the definition of κ as compared to Lemma 9.16.

Take $\Delta_0 < \delta_2$, $\Delta_0 \gtrsim_\eta 1$ sufficiently small so that $\kappa\Delta_0 < \eta/2$. Then for $\delta \in (0, \Delta_0]$ and $|u| < \kappa\delta$, Lemma 9.16 shows

$$\Psi_u(B^n(\delta)) \supseteq B^n(\Psi_u(0), 2\kappa\delta) = B^n(u, 2\kappa\delta) \supseteq B^n(0, \kappa\delta),$$

which completes the proof. $\square$

9.3 Proof of the main result. We turn to the proof of Theorem 4.7. We separate the proof into two parts: when $X_1(x_0), \dots, X_q(x_0)$ are linearly independent (i.e., when $n = q$), and more generally when $X_1(x_0), \dots, X_q(x_0)$ may be linearly dependent (i.e., when $q \geq n$).

9.3.1 Linearly independent. In this section, we prove Theorem 4.7 in the special case $n = q$. We take the same setting as Theorem 4.7 with the same notions of admissible constants, and with the additional assumption that $n = q$. Note that, in this case, $X_{J_0} = X$, so we may replace X_{J_0} with X throughout the statement of Theorem 4.7. Also, because $n = q$, in $\langle m_1, m_2, s \rangle$ -admissible constants, m_2 does not play a role (since in all of our results $m_1 \geq m_2$ when $\langle m_1, m_2, s \rangle$ admissible constants are used), so we instead use $\langle m_1, -1, s \rangle$ -admissible constants throughout this section.

Similarly, we use $\{s, -1\}$ -admissible constants throughout this section.

Proposition 3.1 implies that $B_X(x_0, \xi)$ is an n -dimensional manifold and that $X_1, \dots, X_n$ span the tangent space to every point of $B_X(x_0, \xi)$. Thus, $X_1(y), \dots, X_n(y)$ are linearly independent $\forall y \in B_X(x_0, \xi)$, and Theorem 4.7 (a) follows with $\chi = \xi$. (b) and (c) are both obvious when $n = q$ (and $\chi = \xi$). With (a), (b), and (c) proved, we henceforth assume $c_{j,k}^l \in C_{X_{J_0}}^1(B_{X_{J_0}}(x_0, \xi)) = C_X^1(B_X(x_0, \xi))$, $1 \leq j, k, l \leq n$.

Consider the map $\Phi : B^n(\eta_0) \rightarrow B_X(x_0, \xi)$ defined in (4.1); which we a priori know to be C^1 . Clearly $d\Phi(0) \frac{\partial}{\partial t_j} = X_j(x_0)$. Since $X_1(x_0), \dots, X_n(x_0)$ form a basis of the tangent space $T_{x_0} B_X(x_0, \xi)$, the Inverse Function Theorem shows that there exists a (non-admissible) $\delta > 0$ such that $\Phi : B^n(\delta) \rightarrow \Phi(B^n(\delta))$ is a C^1 diffeomorphism. Let $\hat{Y}_j := \Phi|_{B^n(\delta)}^* X_j$, so that $\hat{Y}_j$ is a C^0 vector field on $B^n(\delta)$. Write $\hat{Y}_j = \frac{\partial}{\partial t_j} + \sum_k \hat{a}_j^k(t) \frac{\partial}{\partial t_k}$. Let $\hat{A}(t) \in C(B^n(\delta); \mathbb{M}^{n \times n})$ denote the $n \times n$ matrix with (j, k) component $\hat{a}_j^k(t)$ and let $C(t) \in C(B^n(\eta_0); \mathbb{M}^{n \times n})$ denote the $n \times n$ matrix with (j, k) component equal to $\sum_{l=1}^n t_l c_{j,l}^k \circ \Phi(t)$.

PROPOSITION 9.18. *Write t in polar coordinates, $t = r\theta$, and consider the differential equation*

$$\frac{\partial}{\partial r} rA(r\theta) = -A(r\theta)^2 - C(r\theta)A(r\theta) - C(r\theta), \quad (9.30)$$

defined for $A : B^n(\eta_0) \rightarrow \mathbb{M}^{n \times n}$. There exists a 0-admissible constant $\eta' > 0$, which also depends on a lower bound for $\eta > 0$, such that there exists a unique continuous

solution $A \in C(B^n(\eta'); \mathbb{M}^{n \times n})$ to (9.30) with $A(0) = 0$. Moreover, this solution lies in $C^1(B^n(\eta'); \mathbb{M}^{n \times n})$ and satisfies

$$\|A(t)\|_{\mathbb{M}^{n \times n}} \lesssim_0 |t| \text{ and } \|A(t)\|_{\mathbb{M}^{n \times n}} \leq \frac{1}{2}, \quad \forall t \in B^n(\eta').$$

For $m \in \mathbb{N}$ and $s \in [0, 1]$, if $c_{i,j}^k \circ \Phi \in C^{m,s}(B^n(\eta'))$ with $\|c_{i,j}^k \circ \Phi\|_{C^{m,s}(B^n(\eta'))} \leq D_{m,s}$, $\forall i, j, k$, then $A \in C^{m,s}(B^n(\eta'); \mathbb{M}^{n \times n})$ and there exists a constant $C_{m,s}$, which depends only on n, m , and $D_{m,s}$, such that $\|A\|_{C^{m,s}(B^n(\eta'); \mathbb{M}^{n \times n})} \leq C_{m,s}$. Similarly, for $s \in (0, \infty)$, if $c_{i,j}^k \circ \Phi \in \mathcal{C}^s(B^n(\eta'))$ with $\|c_{i,j}^k \circ \Phi\|_{\mathcal{C}^s(B^n(\eta'))} \leq D_s$, then there exists a constant C_s which depends only on n, s , and D_s such that $\|A\|_{\mathcal{C}^s(B^n(\eta'); \mathbb{M}^{n \times n})} \leq C_s$. Finally, $\widehat{A}|_{B^n(\min\{\eta', \delta\})} = A|_{B^n(\min\{\eta', \delta\})}$.

Proof. Note that, by the definition of $C(t)$ we have $\|C(t)\|_{\mathbb{M}^{n \times n}} \lesssim_0 |t|$. Also, $\widehat{A}$ satisfies (9.30) on $B^n(\delta)$ by Proposition 9.1. Since $d\Phi(0)\frac{\partial}{\partial t_j} = X_j(x_0)$, we have $\widehat{A}(0) = 0$. With these remarks in hand, the proposition (except for the claim $A \in C^1(B^n(\eta'); \mathbb{M}^{n \times n})$) follows directly from Proposition 9.4 (see also Remark 9.3).

The claim that $A \in C^1(B^n(\eta'); \mathbb{M}^{n \times n})$ can be seen as follows. First note that we may assume $\eta' < \eta_0$ as if $\eta' = \eta_0$, we may replace η' with $\eta_0/2$. Since $c_{j,k}^l \in C_{X_{j_0}}^1(B_{X_{j_0}}(x_0, \xi)) = C_X^1(B_X(x_0, \xi))$, $X_1, \dots, X_n$ span the tangent space at every point of $B_X(x_0, \xi)$, and the vector fields $X_1, \dots, X_n$ are C^1 , it follows that $c_{j,k}^l$ are C^1 on $B_X(x_0, \xi)$. Since $\Phi : B^n(\eta_0) \rightarrow B_X(x_0, \xi)$ is a priori known to be C^1 , we have $c_{j,k}^l \circ \Phi$ is C^1 on $B^n(\eta_0)$. Thus, $C \in C^1(B^n(\eta'); \mathbb{M}^{n \times n})$, and it follows from Proposition 9.4 that $A \in C^1(B^n(\eta'); \mathbb{M}^{n \times n})$. $\square$

We fix $\eta' > 0$ and A as in Proposition 9.18. Write $a_j^k(t)$ for the (j, k) component of $A(t)$ and set $Y_j := \frac{\partial}{\partial t_j} + \sum_{k=1}^n a_j^k \frac{\partial}{\partial t_k}$. Note that $Y_1, \dots, Y_n$ are C^1 vector fields on $B^n(\eta')$. By Proposition 9.18, $Y_j|_{B^n(\min\{\eta', \delta\})} = \widehat{Y}_j|_{B^n(\min\{\eta', \delta\})}$. Since δ is not admissible, we think of δ as being much smaller than η' , and so Y_j should be thought of as extending $\widehat{Y}_j$.

PROPOSITION 9.19. $\forall t \in B^n(\eta')$, $d\Phi(t)Y_j(t) = X_j(\Phi(t))$, $1 \leq j \leq n$.

Proof. Fix $\theta \in S^{n-1}$ and set

$$r_1 := \sup\{r \geq 0 : d\Phi(r'\theta)Y_j(r'\theta) = X_j(\Phi(r'\theta)), 0 \leq r' \leq r, 1 \leq j \leq n\}.$$

We wish to show $r_1 = \eta'$, and this will complete the proof since $\theta \in S^{n-1}$ was arbitrary. Suppose, for contradiction, $r_1 < \eta'$. Since $Y_j|_{B^n(\min\{\eta', \delta\})} = \widehat{Y}_j|_{B^n(\min\{\eta', \delta\})}$ and $d\Phi(u)\widehat{Y}_j(u) = X_j(\Phi(u))$, we know $r_1 > 0$. By continuity, we have

$$d\Phi(r_1\theta)Y_j(r_1\theta) = X_j(\Phi(r_1\theta)).$$

By Proposition 3.1, $X_1(\Phi(r_1\theta)), \dots, X_n(\Phi(r_1\theta))$ span $T_{\Phi(r_1\theta)}B_X(x_0, \xi)$, and therefore the Inverse Function Theorem applies to Φ at the point $r_1\theta$. Thus, there exists

a neighborhood V of $r_1\theta$ such that $\Phi : V \rightarrow \Phi(V)$ is a C^1 diffeomorphism. Pick $0 < r_2 \leq r_3 < r_1 < r_4 < \eta'$ such that $\{r'\theta : r_2 \leq r' \leq r_4\} \subset V$.

Let $\tilde{Y}_j := \Phi|_V^* X_j$. By the choice of r_1 , for $r_2 \leq r' \leq r_3$ we have $\tilde{Y}_j(r'\theta) = Y_j(r'\theta)$. Write $\tilde{Y}_j = \frac{\partial}{\partial u_j} + \sum_{k=1}^n \tilde{a}_j^k \frac{\partial}{\partial u_k}$ and let $\tilde{A}$ denote the matrix with (j, k) component $\tilde{a}_j^k$. We therefore have $\tilde{A}(r'\theta) = A(r'\theta)$ for $r_2 \leq r' \leq r_3$. $\tilde{A}$ satisfies (9.30) by Proposition 9.1. Away from $r = 0$, (9.30) is a standard ODE that both A and $\tilde{A}$ satisfy. Thus, by standard uniqueness theorems (using, for example, Grönwall's inequality) we have $\tilde{A}(r'\theta) = A(r'\theta)$ for $r_2 \leq r' \leq r_4$. Thus, $Y_j(r'\theta) = \tilde{Y}_j(r'\theta)$, $r_2 \leq r' \leq r_4$. Since $d\Phi(r'\theta)\tilde{Y}_j(r'\theta) = X_j(\Phi(r'\theta))$ we conclude $r_1 \geq r_4$. This is a contradiction, completing the proof. $\square$

LEMMA 9.20. $\Phi : B^n(\eta') \rightarrow B_X(x_0, \xi)$ is C^2 .

Proof. Since we already know that $\Phi : B^n(\eta') \rightarrow B_X(x_0, \xi)$ is C^1 , it suffices to show the map $u \mapsto d\Phi(u)$, $u \in B^n(\eta')$ is C^1 . We have already remarked that $Y_1, \dots, Y_n$ are C^1 . Since $Y = (I + A)\nabla$, with $\|A(t)\|_{\mathbb{M}^n \times n} \leq \frac{1}{2}$, $\forall t$, we conclude $Y_1, \dots, Y_n$ are a basis for the tangent space at every point of $B^n(\eta')$. Also, $d\Phi(u)Y_j(u) = X_j(\Phi(u)) \in C^1$ since $X_j \in C^1$, $\Phi \in C^1$. Since $Y_1, \dots, Y_n$ are C^1 and a basis for the tangent space at every point, we conclude $u \mapsto d\Phi(u)$ is C^1 , and therefore Φ is C^2 , completing the proof. $\square$

PROPOSITION 9.21. For $m \in \mathbb{N}$, $s \in [0, 1]$, $\eta'' \in (0, \eta']$ we have (for any function f),

$$\|f\|_{C^{m,s}(B^n(\eta''))} \approx_{\langle m-1, -1, s \rangle} \|f\|_{C_Y^{m,s}(B^n(\eta''))}, \quad (9.31)$$

and

$$\|Y_j\|_{C^{m,s}(B^n(\eta'); \mathbb{R}^n)} \lesssim_{\langle m, -1, s \rangle} 1. \quad (9.32)$$

Similarly, for $s \in (0, \infty)$,

$$\|f\|_{\mathcal{C}^s(B^n(\eta''))} \approx_{\{s-1, -1\}, \eta''} \|f\|_{\mathcal{C}_Y^s(B^n(\eta''))} \quad (9.33)$$

and

$$\|Y_j\|_{\mathcal{C}^s(B^n(\eta'); \mathbb{R}^n)} \lesssim_{\{s, -1\}} 1. \quad (9.34)$$

In (9.33) we have written $\approx_{\{s-1, -1\}, \eta''}$ to denote that the implicit constants are also allowed to depend on the choice of η'' .

Furthermore, for $m \in \mathbb{N}$, $s \in [0, 1]$, and $1 \leq i, j, k \leq n$, we have

$$\|c_{i,j}^k \circ \Phi\|_{C^{m,s}(B^n(\eta'))} \lesssim_{\langle m, -1, s \rangle} 1, \quad (9.35)$$

and for $s \in (0, \infty)$,

$$\|c_{i,j}^k \circ \Phi\|_{\mathcal{C}^s(B^n(\eta'))} \lesssim_{\{s, -1\}} 1. \quad (9.36)$$

Proof. Since $\sup_{t \in B^n(\eta')} \|A(t)\|_{\mathbb{M}^{n \times n}} \leq \frac{1}{2}$, and $Y = (I + A)\nabla$, we also have $\nabla = (I + A)^{-1}Y$. Thus, once we prove a certain regularity on A , we can compare norms as in (9.31) and (9.33) by applying Proposition 8.12. For example, once we show $\|A\|_{C^{m,s}(B^n(\eta'); \mathbb{M}^{n \times n})} \lesssim_{\langle m, -1, s \rangle} 1$, we will also have $\|(I + A)^{-1}\|_{C^{m,s}(B^n(\eta'); \mathbb{M}^{n \times n})} \lesssim_{\langle m, -1, s \rangle} 1$. It will then follow that constants which are $\langle m, s; E \rangle$ -admissible in the sense of Definition 8.9 (when applied to the vector fields $Y_1, \dots, Y_n$) are $\langle m, -1, s \rangle$ -admissible in the sense of Definition 4.3. From here, Proposition 8.12 implies (9.31). Similar comments hold for Zygmund spaces; however, we are applying Proposition 8.12 with η replaced by η'' , and therefore $\{s; E\}$ -admissible constants will also depend on an upper bound for $(\eta'')^{-1}$. This is where the dependance on η'' enters in (9.33).

We first prove (9.31) and (9.32). We claim (for any function f),

$$\|f\|_{C^{m,s}(B^n(\eta''))} \approx_{\langle m-1, -1, s \rangle} \|f\|_{C_Y^{m,s}(B^n(\eta''))}, \quad (9.37)$$

$$\|A\|_{C^{m,s}(B^n(\eta'); \mathbb{M}^{n \times n})} \lesssim_{\langle m, -1, s \rangle} 1, \quad (9.38)$$

which are clearly equivalent to (9.31) and (9.32). We proceed by induction on m . Using that $\|A\|_{C^0(B^n(\eta'); \mathbb{M}^{n \times n})} \leq \frac{1}{2} \lesssim_{\langle -1, -1, s \rangle} 1$, the base case of (9.37) follows from Proposition 8.12. Using this and Propositions 8.6 and 9.19 we have

$$\|c_{i,j}^k \circ \Phi\|_{C^{0,s}(B^n(\eta'))} \approx_{\langle -1, -1, s \rangle} \|c_{i,j}^k \circ \Phi\|_{C_Y^{0,s}(B^n(\eta'))} \leq \|c_{i,j}^k\|_{C_X^{0,s}(B_X(x_0, \xi))} \lesssim_{\langle 0, -1, s \rangle} 1. \quad (9.39)$$

In light of (9.39), Proposition 9.18 implies $\|A\|_{C^{0,s}(B^n(\eta'); \mathbb{M}^{n \times n})} \lesssim_{\langle 0, -1, s \rangle} 1$, completing the proof of the base case $m = 0$.

We assume (9.37) and (9.38) for $m - 1$ and prove them for m . Because $\|A\|_{C^{m-1,s}(B^n(\eta'); \mathbb{M}^{n \times n})} \lesssim_{\langle m-1, -1, s \rangle} 1$, Proposition 8.12 implies (9.37) for m . Thus we need to show (9.38).

Using (9.37) and Propositions 8.6 and 9.19 we have

$$\|c_{i,j}^k \circ \Phi\|_{C^{m,s}(B^n(\eta'))} \approx_{\langle m-1, -1, s \rangle} \|c_{i,j}^k \circ \Phi\|_{C_Y^{m,s}(B^n(\eta'))} \leq \|c_{i,j}^k\|_{C_X^{m,s}(B_X(x_0, \xi))} \lesssim_{\langle m, -1, s \rangle} 1. \quad (9.40)$$

In light of (9.40), Proposition 9.18 implies $\|A\|_{C^{m,s}(B^n(\eta'); \mathbb{M}^{n \times n})} \lesssim_{\langle m, -1, s \rangle} 1$, completing the proof of (9.38), and therefore completing the proof of (9.31) and (9.32).

We turn to proving (9.33) and (9.34). We prove (for any function f)

$$\|f\|_{\mathcal{C}^s(B^n(\eta''))} \approx_{\{s-1, -1\}, \eta''} \|f\|_{\mathcal{C}_Y^s(B^n(\eta''))}, \quad (9.41)$$

$$\|A\|_{\mathcal{C}^s(B^n(\eta'); \mathbb{M}^{n \times n})} \lesssim_{\{s, -1\}} 1, \quad (9.42)$$

which are clearly equivalent to (9.33) and (9.34).

We first prove (9.41) and (9.42) for $s \in (0, 1]$. (9.38) shows

$$\|A\|_{C^{0,s/2}(B^n(\eta'); \mathbb{M}^{n \times n})} \lesssim_{\langle -1, -1, s/2 \rangle} 1,$$

and therefore

$$\|A\|_{C^{0,s/2}(B^n(\eta'); \mathbb{M}^{n \times n})} \lesssim_{\{s-1, -1\}} 1.$$

Using this, Proposition 8.12 implies (9.41). In particular, since η' is a $\{-1, -1\}$ -admissible constant (since it is a 0-admissible constant), and using (9.41) and Propositions 8.6 and 9.19,

$$\|c_{i,j}^k \circ \Phi\|_{\mathcal{C}^s(B^n(\eta'))} \approx_{\{s-1, -1\}} \|c_{i,j}^k \circ \Phi\|_{\mathcal{C}_Y^s(B^n(\eta'))} \leq \|c_{i,j}^k\|_{\mathcal{C}_X^s(B_X(x_0, \xi))} \lesssim_{\{s, -1\}} 1. \quad (9.43)$$

In light of (9.43), Proposition 9.18 implies (9.42).

We now assume (9.41) and (9.42) for $s \in (0, k]$ and prove them for $s \in (k, k+1]$. Fix $s \in (k, k+1]$. By the inductive hypothesis, we know $\|A\|_{\mathcal{C}^{s-1}(B^n(\eta'); \mathbb{M}^{n \times n})} \lesssim_{\{s-1, -1\}} 1$. Using this, Propositions 8.12 implies (9.41) for s . In particular, since η' is a $\{-1, -1\}$ -admissible constant (since it is a 0-admissible constant), and using (9.41) and Propositions 8.6 and 9.19,

$$\|c_{i,j}^k \circ \Phi\|_{\mathcal{C}^s(B^n(\eta'))} \approx_{\{s-1, -1\}} \|c_{i,j}^k \circ \Phi\|_{\mathcal{C}_Y^s(B^n(\eta'))} \leq \|c_{i,j}^k\|_{\mathcal{C}_X^s(B_X(x_0, \xi))} \lesssim_{\{s, -1\}} 1. \quad (9.44)$$

In light of (9.44), Proposition 9.18 implies (9.42).

Finally, (9.35) was established in (9.39) and (9.40) while (9.36) was established in (9.43) and (9.44). $\square$

PROPOSITION 9.22. *There exists a 1-admissible constant $\eta_1 \in (0, \eta']$ such that $\Phi|_{B^n(\eta_1)}$ is injective. Furthermore, $\Phi(B^n(\eta_1)) \subseteq B_X(x_0, \xi)$ is open and $\Phi : B^n(\eta_1) \rightarrow \Phi(B^n(\eta_1))$ is a C^2 -diffeomorphism.*

Proof. Consider the maps, defined for $u, v \in \mathbb{R}^n$ sufficiently small, given by

$$\Psi_u(v) = e^{v_1 Y_1 + \dots + v_n Y_n} u.$$

Since $Y = (I + A)\nabla$ and $\|A(t)\|_{\mathbb{M}^{n \times n}} \leq \frac{1}{2}$, $\forall t \in B^n(\eta')$, we have $|\det(Y_1(t)) \cdots Y_n(t)| \geq c_n > 0$, $\forall t \in B^n(\eta')$, where $c_n > 0$ can be chosen to depend only on n . Furthermore, by Proposition 9.21 (taking $m = 1$, $s = 0$ in (9.32)), we have

$$\|Y_j\|_{C^1(B^n(\eta'); \mathbb{R}^n)} \lesssim_{\langle 1, -1, 0 \rangle} 1. \quad (9.45)$$

Thus, by the definition of 1-admissible constants, we have $\|Y_j\|_{C^1(B^n(\eta'); \mathbb{R}^n)} \lesssim_1 1$.

Take $\Delta_0, \kappa > 0$ as in Proposition 9.15 (with η' playing the role of η in that proposition). In light of the above remarks, Δ_0 and κ can be taken to be 1-admissible constants. Set $\delta_1 := \min\{\Delta_0, \delta_0, 1\}$ so that $\delta_1 > 0$ is a 1-admissible constant; see Section 3.2 for the definition of δ_0 . Let $\eta_1 := \min\{\delta_1 \kappa, \eta'\} > 0$ so that η_1 is a 1-admissible constant.

We claim $\Phi|_{B^n(\eta_1)}$ is injective. Let $u_1, u_2 \in B^n(\eta_1)$ be such that $\Phi(u_1) = \Phi(u_2)$; we wish to show $u_1 = u_2$. By Proposition 9.15 there exists $v \in B^n(\delta_1)$ such that $u_2 = \Psi_{u_1}(v)$, i.e., $u_2 = e^{v \cdot Y} u_1$. Since $d\Phi(u) Y_j(u) = X_j(\Phi(u))$ (Proposition 9.19), it follows that

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

Also, we know $X_1(\Phi(u)), \dots, X_n(\Phi(u))$ are linearly independent (as a consequence of Proposition 3.1). Finally, X satisfies $\mathcal{C}(\Phi(u_1), \delta_1, B_X(x_0, \xi))$ because Y satisfies

$\mathcal{C}(u_1, \delta_1, B^n(\eta'))$ (by Proposition 9.15). Hence, by the definition of δ_0 , we have $v = 0$. We conclude $u_2 = e^{v \cdot Y} u_1 = u_1$, and therefore Φ is injective.

Combining the fact that $d\Phi(u)Y_j(u) = X_j(\Phi(u))$ and $X_1, \dots, X_n$ span the tangent space at every point of $B_X(x_0, \xi)$, the Inverse Function Theorem implies $\Phi : B^n(\eta') \rightarrow B_X(x_0, \xi)$ is an open map and is locally a C^1 diffeomorphism. In particular, $\Phi(B^n(\eta_1))$ is open. Hence, since Φ is injective, locally a C^1 diffeomorphism, and Φ is C^2 (Lemma 9.20), we conclude $\Phi : B^n(\eta_1) \rightarrow \Phi(B^n(\eta_1))$ is a C^2 -diffeomorphism. $\square$

LEMMA 9.23. *There exists a 1-admissible constant $\xi_1 > 0$ such that $B_X(x_0, \xi_1) \subseteq \Phi(B^n(\eta_1))$.*

Proof. Fix $\xi_1 \in (0, \xi]$ to be chosen later, and suppose $y \in B_X(x_0, \xi_1)$. Thus, there exists $\gamma : [0, 1] \rightarrow B_X(x_0, \xi)$ with $\gamma(0) = x_0$, $\gamma(1) = y$, $\gamma'(t) = \sum_{j=1}^n b_j(t) \xi_1 X_j(\gamma(t))$, $\|\sum |b_j(t)|^2\|_{L^\infty([0,1])} < 1$. Define

$$t_0 := \sup\{t \in [0, 1] : \gamma(t') \in \Phi(B^n(\eta_1/2)), \forall 0 \leq t' \leq t\}.$$

We want to show that by taking $\xi_1 > 0$ to be a sufficiently small 1-admissible constant, we have $t_0 = 1$ and $\gamma(1) \in \Phi(B^n(\eta_1/2))$. Note that $t_0 \geq 0$, since $\gamma(0) = x_0 = \Phi(0)$.

Suppose not. Then $|\Phi^{-1}(\gamma(t_0))| = \frac{\eta_1}{2}$. And, using that $\|Y_j\|_{C(B^n(\eta_1); \mathbb{R}^n)} \lesssim_0 1$ and $\Phi(0) = x_0$,

$$\begin{aligned} \eta_1/2 = |\Phi^{-1}(\gamma(t_0))| &= \left| \int_0^{t_0} \frac{d}{dt} \Phi^{-1} \circ \gamma(t) dt \right| = \left| \int_0^{t_0} \sum_{j=1}^n b_j(t) \xi_1 Y_j(\Phi^{-1} \circ \gamma(t)) dt \right| \\ &\lesssim_0 \xi_1. \end{aligned}$$

This is a contradiction if ξ_1 is a sufficiently small 1-admissible constant, completing the proof. $\square$

LEMMA 9.24. $[Y_i, Y_j] = \sum_{k=1}^n \tilde{c}_{i,j}^k Y_k$ on $B^n(\eta_1)$, where $\|\tilde{c}_{i,j}^k\|_{C^{m,s}(B^n(\eta_1))} \lesssim_{\langle m, -1, s \rangle} 1$ and $\|\tilde{c}_{i,j}^k\|_{\mathcal{C}^s(B^n(\eta_1))} \lesssim_{\{s, -1\}} 1$.

Proof. Because $\Phi : B^n(\eta_1) \rightarrow \Phi(B^n(\eta_1))$ is a diffeomorphism, we have

$$[Y_i, Y_j] = [\Phi^* X_i, \Phi^* X_j] = \Phi^* [X_i, X_j] = \Phi^* \sum_k c_{i,j}^k X_k = \sum_k \tilde{c}_{i,j}^k Y_k,$$

with $\tilde{c}_{i,j}^k = c_{i,j}^k \circ \Phi$. From here the result follows from Proposition 9.21, since $\eta_1 \leq \eta'$. $\square$

Proof of Theorem 4.7 when $n = q$. As mentioned above, we take $\chi := \xi$. We also take $\xi_2 := \xi_1$. Note that (i) is vacuous when $n = q$. Also, since $n = q$, $X = X_{J_0}$ and $Y = Y_{J_0}$. With these remarks, all of the parts of Theorem 4.7 except for (l) were proved above. We clarify one point in (k). In Proposition 9.21, (k) was proved

on $B^n(\eta'')$ for any $\eta'' \in (0, \eta']$. Here, we are taking $\eta'' = \eta_1$. However, in the case of Zygmund spaces the implicit constant in (9.33) also depended on the choice of η'' . Since η_1 is a 1-admissible constant, if $s > 2$, it is a $\{s-1, -1\}$ -admissible constant. This is why (k) is only stated for $s > 2$ in the case of Zygmund spaces—in the case $s \leq 1$, the implicit constants also depend on η_1 , and are therefore 1-admissible constants.¹⁵

We close the proof by proving (l). We prove the result for Zygmund spaces, the same proof works for Hölder spaces. Let $f \in C(B_{X_{J_0}}(x_0, \chi))$. We use Proposition 9.21 in the case $\eta'' = \eta'$, and that η' is a $\{-1, -1\}$ -admissible constant. We also use Proposition 8.6. We have, for $s \in (0, \infty)$,

$$\|f \circ \Phi\|_{\mathcal{C}^s(B^n(\eta_1))} \leq \|f \circ \Phi\|_{\mathcal{C}^s(B^n(\eta'))} \approx_{\{s-1, -1\}} \|f \circ \Phi\|_{\mathcal{C}_V^s(B^n(\eta'))} \leq \|f\|_{\mathcal{C}_X^s(B_X(x_0, \chi))},$$

completing the proof. $\square$

In the third paper of this series, it will be convenient to use a slight modification of Theorem 4.7 in the case $n = q$, where we replace 1-admissible constants with a slightly different definition. We present this here.

DEFINITION 9.25. *In the case $n = q$, if we say C is a $1'$ -admissible constant, it means that we assume $c_{j,k}^l \circ \Phi \in C^1(B^n(\eta_0))$, for $1 \leq j, k, l \leq n$. C is then allowed to depend only on upper bounds for n , ξ^{-1} , η^{-1} , δ_0^{-1} , and $\|c_{j,k}^l \circ \Phi\|_{C^1(B^n(\eta_0))}$ and $\|c_{j,k}^l\|_{C(B_{X_{J_0}}(x_0, \xi))}$ ($1 \leq j, k, l \leq n$).*

PROPOSITION 9.26. *In the case $n = q$, Theorem 4.7 (except for (k)) holds with the following modifications. The assumption $c_{j,k}^l \in C_{X_{J_0}}^1(B_{X_{J_0}}(x_0, \xi))$ is replaced by $c_{j,k}^l \circ \Phi \in C^1(B^n(\eta_0))$ and 1-admissible constants are replaced with $1'$ -admissible constants throughout.*

Comments on the proof. The only place the estimates on $\|c_{j,k}^l\|_{C_{X_{J_0}}^1(B_{X_{J_0}}(x_0, \xi))}$ from 1-admissible constants arose in the proof was to conclude $\|Y_j\|_{C^1(B^n(\eta'); \mathbb{R}^n)} \lesssim_1 1$; i.e., to conclude $\|A\|_{C^1(B^n(\eta'); \mathbb{M}^{n \times n})} \lesssim_1 1$. However, one obtains $\|A\|_{C^1(B^n(\eta'); \mathbb{M}^{n \times n})} \lesssim_{1'} 1$ directly from Proposition 9.18. Using this, the proof goes through unchanged. $\square$

9.3.2 Linearly dependent. In this section, we prove Theorem 4.7 in the general case $q \geq n$. Thus, we take the same setting and notation as in Theorem 4.7.

LEMMA 9.27. *For $J \in \mathcal{I}(n, q)$, $1 \leq j \leq n$,*

$$\mathcal{L}_{X_j} \bigwedge X_J = \sum_{K \in \mathcal{I}_0(n, q)} g_{j,J}^K \bigwedge X_K, \text{ on } B_{X_{J_0}}(x_0, \xi),$$

¹⁵ It is classical that $C^{0,s}(B^n(\eta_1))$ and $\mathcal{C}^s(B^n(\eta_1))$ have comparable norms for $s \in (0, 1)$. However, the constants involved in the comparability of these norms depend on η_1 , and are therefore 1-admissible.

where

$$\|g_{j,J}^K\|_{C(B_{X_{J_0}}(x_0,\xi))} \lesssim_0 1,$$

for $m \in \mathbb{N}$ and $s \in [0, 1]$,

$$\|g_{j,J}^K\|_{C_X^{m,s}(B_{X_{J_0}}(x_0,\xi))} \lesssim_{\langle m,m,s \rangle} 1,$$

and for $s \in (0, \infty)$,

$$\|g_{j,J}^K\|_{\mathcal{C}_X^s(B_{X_{J_0}}(x_0,\xi))} \lesssim_{\{s,s\}} 1.$$

Proof. Let $J = (j_1, \dots, j_n)$. We have,

$$\begin{aligned} \mathcal{L}_{X_j} \bigwedge X_J &= \mathcal{L}_{X_j} (X_{j_1} \wedge X_{j_2} \wedge \dots \wedge X_{j_n}) \\ &= \sum_{l=1}^n X_{j_1} \wedge X_{j_2} \wedge \dots \wedge X_{j_{l-1}} \wedge [X_j, X_{j_l}] \wedge X_{j_{l+1}} \wedge \dots \wedge X_{j_n} \\ &= \sum_{l=1}^n \sum_{k=1}^q c_{j,j_l}^k X_{j_1} \wedge X_{j_2} \wedge \dots \wedge X_{j_{l-1}} \wedge X_k \wedge X_{j_{l+1}} \wedge \dots \wedge X_{j_n}. \end{aligned}$$

The result follows from the anti-commutativity of $\wedge$ and the assumptions on $c_{i,j}^k$. $\square$

LEMMA 9.28. Let $\chi' \in (0, \xi]$. Suppose for all $y \in B_{X_{J_0}}(x_0, \chi')$, $\bigwedge X_{J_0}(y) \neq 0$. Then, for $J \in \mathcal{I}(n, q)$, $1 \leq j \leq n$,

$$X_j \frac{\bigwedge X_J}{\bigwedge X_{J_0}} = \sum_{K \in \mathcal{I}_0(n, q)} g_{j,J}^K \frac{\bigwedge X_K}{\bigwedge X_{J_0}} - \sum_{K \in \mathcal{I}_0(n, q)} g_{j,J_0}^K \frac{\bigwedge X_J}{\bigwedge X_{J_0}} \frac{\bigwedge X_K}{\bigwedge X_{J_0}} \text{ on } B_{X_{J_0}}(x_0, \chi'),$$

where $g_{j,J}^K$ are the functions from Lemma 9.27.

Proof. This follows by combining Lemmas 5.1 and 9.27. $\square$

LEMMA 9.29. Let $C > 0$ and $u_0 > 0$. Let $u_{u_0, C}(t)$ be the unique solution to

$$\frac{d}{dt} u_{u_0, C}(t) = C(u_{u_0, C}(t) + u_{u_0, C}(t)^2), \quad u_{u_0, C}(0) = u_0,$$

defined on some maximum interval $[0, R_{u_0, C})$. Let $F(t)$ be a non-negative function defined on $[0, R')$ with $R' \leq R_{u_0, C}$ satisfying

$$\frac{d}{dt} F(t) \leq C(F(t) + F(t)^2), \quad F(0) \leq u_0.$$

Then, for $t \in [0, R')$, $F(t) \leq u_{u_0, C}(t)$.

Proof. This is standard and is easy to see directly. It is also a special case of the Bihari–LaSalle inequality. $\square$

LEMMA 9.30. *There exists a 0-admissible constant $\chi \in (0, \xi]$ such that the following holds. Suppose $\gamma : [0, \chi] \rightarrow B_{X_{J_0}}(x_0, \xi)$ satisfies $\gamma(0) = x_0$, $\gamma'(t) = \sum_{j=1}^n a_j(t) X_j(\gamma(t))$, and $\|\sum |a_j(t)|^2\|_{L^\infty([0, \chi])} < 1$. Suppose further that for some $\chi' \in (0, \chi]$, $\bigwedge X_{J_0}(\gamma(t)) \neq 0$ for $t \in (0, \chi']$. Then,*

$$\sup_{\substack{J \in \mathcal{I}(n, q) \\ t \in [0, \chi']}} \left| \frac{\bigwedge X_J(\gamma(t))}{\bigwedge X_{J_0}(\gamma(t))} \right| \lesssim_0 1. \quad (9.46)$$

Here, the implicit constant depends on neither χ' nor γ .

Proof. Let $\chi \in (0, \xi]$ be a 0-admissible constant to be chosen later. Let γ and χ' be as in the statement of the lemma. We wish to show that if χ is chosen to be a sufficiently small 0-admissible constant (which forces χ' to be small), then (9.46) holds.

Set

$$F(t) := \sum_{J \in \mathcal{I}_0(n, q)} \left| \frac{X_J(\gamma(t))}{X_{J_0}(\gamma(t))} \right|^2.$$

We wish to show that if χ is a sufficiently small 0-admissible constant, then $F(t) \lesssim_0 1$, $\forall t \in [0, \chi']$, and this will complete the proof.¹⁶

Using Lemma 9.28, we have,

$$\begin{aligned} \frac{d}{dt} F(t) &= \sum_{J \in \mathcal{I}_0(n, q)} 2 \frac{\bigwedge X_J(\gamma(t))}{\bigwedge X_{J_0}(\gamma(t))} \sum_{j=1}^n a_j(t) \left(X_j \frac{\bigwedge X_J}{\bigwedge X_{J_0}} \right) (\gamma(t)) \\ &= \sum_{J \in \mathcal{I}_0(n, q)} \sum_{K \in \mathcal{I}_0(n, q)} \sum_{j=1}^n 2a_j(t) \frac{\bigwedge X_J(\gamma(t))}{\bigwedge X_{J_0}(\gamma(t))} \left(g_{j, J}^K(\gamma(t)) \frac{\bigwedge X_K(\gamma(t))}{\bigwedge X_{J_0}(\gamma(t))} - g_{j, J_0}^K(\gamma(t)) \frac{\bigwedge X_J(\gamma(t))}{\bigwedge X_{J_0}(\gamma(t))} \frac{\bigwedge X_K(\gamma(t))}{\bigwedge X_{J_0}(\gamma(t))} \right) \\ &\lesssim_0 F(t) + F(t)^{3/2} \lesssim_0 F(t) + F(t)^2. \end{aligned}$$

Also, we have

$$F(0) = \sum_{J \in \mathcal{I}_0(n, q)} \left| \frac{X_J(x_0)}{X_{J_0}(x_0)} \right|^2 \lesssim_0 1.$$

Thus, there exist 0-admissible constants C and $u_0 > 0$ such that

$$\frac{d}{dt} F(t) \leq C (F(t) + F(t)^2), \quad F(0) \leq u_0.$$

Standard theorems from ODEs show that if $\chi = \chi(C, u_0) > 0$ is chosen sufficiently small, then the unique solution $u(t)$ to

$$\frac{d}{dt} u(t) = C (u(t) + u(t)^2), \quad u(0) = u_0,$$

¹⁶ Here we are using $\forall K \in \mathcal{I}(n, q)$, either $\bigwedge X_K \equiv 0$ or $\exists J \in \mathcal{I}_0(n, q)$ with $\bigwedge X_K = \pm \bigwedge X_J$, by the basic properties of wedge products.

exists for $t \in [0, \chi]$ and satisfies $u(t) \leq 2u_0$, $\forall t \in [0, \chi]$. For this choice of χ (which is 0-admissible, since C and u_0 are), Lemma 9.29 shows $F(t) \leq 2u_0 \lesssim_0 1$, $\forall t \in [0, \chi']$, completing the proof. $\square$

PROPOSITION 9.31. *There exists a 0-admissible constant $\chi \in (0, \xi]$ such that $\forall y \in B_{X_{J_0}}(x_0, \chi)$, $\bigwedge X_{J_0}(y) \neq 0$ and*

$$\sup_{\substack{J \in \mathcal{I}(n, q) \\ y \in B_{X_{J_0}}(x_0, \chi)}} \left| \frac{\bigwedge X_J(y)}{\bigwedge X_{J_0}(y)} \right| \lesssim_0 1. \quad (9.47)$$

Proof. Take χ as in Lemma 9.30. First we claim $\forall y \in B_{X_{J_0}}(x_0, \chi)$, $\bigwedge X_{J_0}(y) \neq 0$. Fix $y \in B_{X_{J_0}}(x_0, \chi)$, so that there exists $\gamma : [0, \chi] \rightarrow B_{X_{J_0}}(x_0, \xi)$, $\gamma(0) = x_0$, $\gamma(\chi) = y$, $\gamma'(t) = \sum_{j=1}^n a_j(t) X_j(\gamma(t))$, $\|\sum |a_j(t)|^2\|_{L^\infty([0,1])} < 1$. We will show that $\forall t \in [0, \chi]$, $\bigwedge X_{J_0}(\gamma(t)) \neq 0$, and then it will follow that $\bigwedge X_{J_0}(y) = \bigwedge X_{J_0}(\gamma(\chi)) \neq 0$.

Suppose not, so that $\bigwedge X_{J_0}(\gamma(t)) = 0$ for some $t \in [0, \chi]$. Let $t_0 = \inf\{t \in [0, \chi] : \bigwedge X_{J_0}(\gamma(t)) = 0\}$, so that $\bigwedge X_{J_0}(\gamma(t_0)) = 0$ but $\bigwedge X_{J_0}(\gamma(t)) \neq 0$, $\forall t \in [0, t_0)$. Note that $t_0 > 0$ since $\bigwedge X_{J_0}(x_0) \neq 0$.

Let ν be a C^1 n -form, defined on a neighborhood of $\gamma(t_0)$ and which is nonzero at $\gamma(t_0)$. We have

$$\lim_{t \uparrow t_0} \nu \left(\bigwedge X_{J_0} \right) (\gamma(t)) = 0, \quad \lim_{t \uparrow t_0} \max_{J \in \mathcal{I}(n, q)} |\nu(X_J)(\gamma(t))| > 0,$$

by continuity, the fact that $X_1, \dots, X_q$ span the tangent space at $\gamma(t_0)$, and that ν is nonzero at $\gamma(t_0)$. We conclude,

$$\lim_{t \uparrow t_0} \sup_{J \in \mathcal{I}(n, q)} \left| \frac{\bigwedge X_J(\gamma(t))}{\bigwedge X_{J_0}(\gamma(t))} \right| = \lim_{t \uparrow t_0} \sup_{J \in \mathcal{I}(n, q)} \left| \frac{\nu(\bigwedge X_J)(\gamma(t))}{\nu(\bigwedge X_{J_0})(\gamma(t))} \right| = \infty. \quad (9.48)$$

Take any $\chi' \in (0, t_0)$. We know $\forall t \in [0, \chi']$, $\bigwedge X_{J_0}(\gamma(t)) \neq 0$. Lemma 9.30 implies

$$\sup_{\substack{J \in \mathcal{I}(n, q) \\ t \in [0, \chi']}} \left| \frac{\bigwedge X_J(\gamma(t))}{\bigwedge X_{J_0}(\gamma(t))} \right| \lesssim_0 1.$$

Since $\chi' \in (0, t_0)$ was arbitrary, we have

$$\sup_{\substack{J \in \mathcal{I}(n, q) \\ t \in [0, t_0)}} \left| \frac{\bigwedge X_J(\gamma(t))}{\bigwedge X_{J_0}(\gamma(t))} \right| \lesssim_0 1.$$

This contradicts (9.48) and completes the proof that $\bigwedge X_{J_0}(y) \neq 0$, $\forall y \in B_{X_{J_0}}(x_0, \chi)$.

To prove (9.47) take $y \in B_{X_{J_0}}(x_0, \chi)$. Then, there exists $\gamma : [0, \chi] \rightarrow B_{X_{J_0}}(x_0, \xi)$, $\gamma(0) = x_0$, $\gamma(\chi) = y$, $\gamma'(t) = \sum_{j=1}^n a_j(t) X_j(\gamma(t))$, $\|\sum |a_j(t)|^2\|_{L^\infty([0, \chi])} < 1$.

We have already shown $\bigwedge X_{J_0}(\gamma(t)) \neq 0$, $\forall t \in (0, \chi]$. Lemma 9.30 implies $\sup_{J \in \mathcal{I}(n,q)} \left| \frac{\bigwedge X_J(y)}{\bigwedge X_{J_0}(y)} \right| = \sup_{J \in \mathcal{I}(n,q)} \left| \frac{\bigwedge X_J(\gamma(\chi))}{\bigwedge X_{J_0}(\gamma(\chi))} \right| \lesssim_0 1$. Since $y \in B_{X_{J_0}}(x_0, \chi)$ was arbitrary, (9.47) follows. $\square$

For the remainder of the section, fix $\chi \in (0, \xi]$ as in Proposition 9.31.

LEMMA 9.32. For $m \in \mathbb{N}$, $s \in [0, 1]$, $J \in \mathcal{I}(n, q)$,

$$\left\| \frac{\bigwedge X_J}{\bigwedge X_{J_0}} \right\|_{C_{X_{J_0}}^{m,s}(B_{X_{J_0}}(x_0, \chi))} \lesssim_{\langle m-1, m-1, s \rangle} 1, \quad (9.49)$$

and for $s \in (0, \infty)$,

$$\left\| \frac{\bigwedge X_J}{\bigwedge X_{J_0}} \right\|_{\mathcal{C}_{X_{J_0}}^s(B_{X_{J_0}}(x_0, \chi))} \lesssim_{\{s-1, s-1\}} 1. \quad (9.50)$$

Proof. In this proof, we freely use the estimates on the functions $g_{j,J}^K$ as described in Lemmas 9.27 and 9.28. We begin with (9.49). Proposition 9.31 shows

$$\left\| \frac{\bigwedge X_J}{\bigwedge X_{J_0}} \right\|_{C(B_{X_{J_0}}(x_0, \chi))} \lesssim_0 1. \quad (9.51)$$

We claim,

$$\left\| \frac{\bigwedge X_J}{\bigwedge X_{J_0}} \right\|_{C_{X_{J_0}}^1(B_{X_{J_0}}(x_0, \chi))} \lesssim_0 1. \quad (9.52)$$

Indeed, for $1 \leq j \leq n$, using Lemma 9.28,

$$\begin{aligned} & \left\| X_j \frac{\bigwedge X_J}{\bigwedge X_{J_0}} \right\|_{C(B_{X_{J_0}}(x_0, \chi))} \\ &= \left\| \sum_{K \in \mathcal{I}_0(n,q)} g_{j,J}^K \frac{\bigwedge X_K}{\bigwedge X_{J_0}} - \sum_{K \in \mathcal{I}_0(n,q)} g_{j,J_0}^K \frac{\bigwedge X_J}{\bigwedge X_{J_0}} \frac{\bigwedge X_K}{\bigwedge X_{J_0}} \right\|_{C(B_{X_{J_0}}(x_0, \chi))} \\ &\lesssim_0 \sum_{K \in \mathcal{I}_0(n,q)} \left\| \frac{\bigwedge X_K}{\bigwedge X_{J_0}} \right\|_{C(B_{X_{J_0}}(x_0, \chi))} + \left\| \frac{\bigwedge X_J}{\bigwedge X_{J_0}} \right\|_{C(B_{X_{J_0}}(x_0, \chi))} \left\| \frac{\bigwedge X_K}{\bigwedge X_{J_0}} \right\|_{C(B_{X_{J_0}}(x_0, \chi))} \lesssim_0 1, \end{aligned}$$

where the last inequality follows from (9.51). (9.52) follows.

Using Lemma 8.1 (i) and (ii), we have for $s \in [0, 1]$,

$$\begin{aligned} \left\| \frac{\bigwedge X_J}{\bigwedge X_{J_0}} \right\|_{C_{X_{J_0}}^{0,s}(B_{X_{J_0}}(x_0, \chi))} &\leq 3 \left\| \frac{\bigwedge X_J}{\bigwedge X_{J_0}} \right\|_{C_{X_{J_0}}^{0,1}(B_{X_{J_0}}(x_0, \chi))} \\ &\leq 3 \left\| \frac{\bigwedge X_J}{\bigwedge X_{J_0}} \right\|_{C_{X_{J_0}}^1(B_{X_{J_0}}(x_0, \chi))} \lesssim_0 1, \end{aligned}$$

where the last inequality used (9.52). This proves (9.49) in the case $m = 0$.

We prove (9.49) by induction on m , the base case ($m = 0$) having just been proved. We assume (9.49) for $m - 1$ and prove it for m . We use Proposition 8.3 freely in what follows. We have

$$\begin{aligned} & \left\| \frac{\bigwedge X_J}{\bigwedge X_{J_0}} \right\|_{C_{X_{J_0}}^{m,s}(B_{X_{J_0}}(x_0, \chi))} \\ &= \left\| \frac{\bigwedge X_J}{\bigwedge X_{J_0}} \right\|_{C_{X_{J_0}}^{m-1,s}(B_{X_{J_0}}(x_0, \chi))} + \sum_{j=1}^n \left\| X_j \frac{\bigwedge X_J}{\bigwedge X_{J_0}} \right\|_{C_{X_{J_0}}^{m-1,s}(B_{X_{J_0}}(x_0, \chi))}. \end{aligned}$$

The first term is $\lesssim_{\langle m-2, m-2, s \rangle} 1$, by the inductive hypothesis, so we focus only on the second term. We have, using Lemma 9.28, and letting C_m be a constant which depends only on m ,

$$\begin{aligned} & \left\| X_j \frac{\bigwedge X_J}{\bigwedge X_{J_0}} \right\|_{C_{X_{J_0}}^{m-1,s}(B_{X_{J_0}}(x_0, \chi))} \\ & \leq C_m \sum_{K \in \mathcal{I}_0(n, q)} \|g_{j,J}^K\|_{C_{X_{J_0}}^{m-1,s}(B_{X_{J_0}}(x_0, \xi))} \left\| \frac{\bigwedge X_K}{\bigwedge X_{J_0}} \right\|_{C_{X_{J_0}}^{m-1,s}(B_{X_{J_0}}(x_0, \chi))} \\ & + C_m \sum_{K \in \mathcal{I}_0(n, q)} \|g_{j,J_0}^K\|_{C_{X_{J_0}}^{m-1,s}(B_{X_{J_0}}(x_0, \xi))} \left\| \frac{\bigwedge X_J}{\bigwedge X_{J_0}} \right\|_{C_{X_{J_0}}^{m-1,s}(B_{X_{J_0}}(x_0, \chi))} \left\| \frac{\bigwedge X_K}{\bigwedge X_{J_0}} \right\|_{C_{X_{J_0}}^{m-1,s}(B_{X_{J_0}}(x_0, \chi))} \\ & \lesssim_{\langle m-1, m-1, s \rangle} 1, \end{aligned}$$

where the last inequality follows from the bounds described in Lemma 9.27 and the inductive hypothesis. This completes the proof of (9.49).

We turn to (9.50), and proceed by induction on m , where $s \in (m, m+1]$. We begin with the base case, $m = 0$, so that $s \in (0, 1]$. Using Lemma 8.1 (iii), we have

$$\left\| \frac{\bigwedge X_J}{\bigwedge X_{J_0}} \right\|_{\mathcal{C}_{X_{J_0}}^s(B_{X_{J_0}}(x_0, \chi))} \leq 5 \left\| \frac{\bigwedge X_J}{\bigwedge X_{J_0}} \right\|_{C_{X_{J_0}}^{0,s}(B_{X_{J_0}}(x_0, \chi))} \lesssim_0 1,$$

where the last inequality follows from (9.49). This implies (9.50) for the base case $s \in (0, 1]$. From here, the inductive step follows just as in (9.49) and we leave the remaining details to the reader. $\square$

LEMMA 9.33. *For $1 \leq k \leq q$, $1 \leq l \leq n$, there exists $\tilde{b}_k^l \in C(B_{X_{J_0}}(x_0, \chi))$ such that*

$$X_k = \sum_{l=1}^n \tilde{b}_k^l X_l, \quad (9.53)$$

where for $m \in \mathbb{N}$, $s \in [0, 1]$,

$$\left\| \tilde{b}_k^l \right\|_{C_{X_{J_0}}^{m,s}(B_{X_{J_0}}(x_0, \chi))} \lesssim_{\langle m-1, m-1, s \rangle} 1,$$

and for $s \in (0, \infty)$,

$$\left\| \tilde{b}_k^l \right\|_{\mathcal{C}_{X_{J_0}}^s(B_{X_{J_0}}(x_0, \chi))} \lesssim_{\{s-1, s-1\}} 1.$$

Proof. For $1 \leq k \leq n$ this is trivial (merely take $\tilde{b}_k^l = 1$ if $k = l$ and $\tilde{b}_k^l = 0$ if $k \neq l$), however the proof that follows deals with all $1 \leq k \leq q$ simultaneously.

For $1 \leq k \leq q$, $1 \leq l \leq n$, let $J(l, k) = (1, 2, \dots, l-1, k, l+1, \dots, n) \in \mathcal{I}(n, q)$. We have, by Cramer's rule (5.1),

$$X_k = \sum_{l=1}^n \frac{\bigwedge X_{J(l,k)}}{\bigwedge X_{J_0}} X_l.$$

From here, the result follows from Lemma 9.32. $\square$

LEMMA 9.34. For $1 \leq i, j, l \leq n$, $\exists \hat{c}_{i,j}^l \in C(B_{X_{J_0}}(x_0, \chi))$ such that $[X_i, X_j] = \sum_{l=1}^n \hat{c}_{i,j}^l X_l$, where for $m \in \mathbb{N}$, $s \in [0, 1]$,

$$\|\hat{c}_{i,j}^l\|_{C_{X_{J_0}}^{m,s}(B_{X_{J_0}}(x_0, \chi))} \lesssim_{\langle m, m-1, s \rangle} 1,$$

and for $s \in (0, \infty)$,

$$\|\hat{c}_{i,j}^l\|_{\mathcal{C}_{X_{J_0}}^s(B_{X_{J_0}}(x_0, \chi))} \lesssim_{\{s, s-1\}} 1.$$

Proof. For $1 \leq i, j \leq n$ and using Lemma 9.33, we have

$$[X_i, X_j] = \sum_{k=1}^q c_{i,j}^k X_k = \sum_{l=1}^n \left(\sum_{k=1}^q c_{i,j}^k \tilde{b}_k^l \right) X_l.$$

Setting $\hat{c}_{i,j}^l = \sum_{k=1}^q c_{i,j}^k \tilde{b}_k^l$, the result follows from the definition of admissible constants, Lemma 9.33, and Proposition 8.3. $\square$

Lemma 9.34 shows that the case $n = q$ of Theorem 4.7 (which was proved in Section 9.3.1) applies to $X_1, \dots, X_n$, with ξ replaced by χ .¹⁷ In light of Lemma 9.34 any constants which are $\langle m, m-1, s \rangle$, $\{s, s-1\}$, 0, or 1-admissible in the sense of this application of the case $n = q$ of Theorem 4.7, are $\langle m, m-1, s \rangle$, $\{s, s-1\}$, 0, or 1-admissible (respectively) in the sense of this section. Thus, from the case $n = q$, we obtain 1-admissible constants $\xi_1, \eta_1 > 0$ and a map $\Phi : B^n(\eta_1) \rightarrow B_{X_{J_0}}(x_0, \chi)$ as in Theorem 4.7. Most of the case $q \geq n$ of Theorem 4.7 immediately follows from this application of the case $n = q$. All that remain to show are: (b), (c), there exists ξ_2 as in (f), (h) for $n+1 \leq j \leq q$, (i), and (k).

Proof of (b). That $\forall y \in B_{X_{J_0}}(x_0, \chi)$

$$\sup_{J \in \mathcal{I}(n, q)} \left| \frac{\bigwedge X_J(y)}{\bigwedge X_{J_0}(y)} \right| \geq 1$$

is clear (by taking $J = J_0$). That

$$\sup_{J \in \mathcal{I}(n, q)} \left| \frac{\bigwedge X_J(y)}{\bigwedge X_{J_0}(y)} \right| \lesssim_0 1,$$

$\forall y \in B_{X_{J_0}}(x_0, \chi)$, is Proposition 9.31. $\square$

¹⁷ When we proved Theorem 4.7 for $n = q$, in Section 9.3.1, we took $\chi = \xi$.

Proof of (c). Let $\chi' \in (0, \chi]$ and fix $x \in B_{X_{j_0}}(x_0, \chi)$. (a) shows $X_1(x), \dots, X_n(x)$ are linearly independent. Define $\Psi(t) := e^{t_1 X_1 + \dots + t_n X_n} x$, so that $d\Psi(0) = (X_1(x) | \dots | X_n(x))$ and is therefore invertible. It is clear that for δ sufficiently small $\Psi(B^n(\delta)) \subseteq B_{X_{j_0}}(x_0, \chi')$ and the Inverse Function Theorem shows that for δ sufficiently small $\Psi(B^n(\delta)) \subseteq B_X(x_0, \xi)$ is open. Hence, $\Psi(B^n(\delta))$ is an open neighborhood of x in $B_{X_{j_0}}(x_0, \chi')$. Since $x \in B_{X_{j_0}}(x_0, \chi')$ was arbitrary, we conclude $B_{X_{j_0}}(x_0, \chi') \subseteq B_X(x_0, \xi)$ is open. $\square$

That there exists a 1-admissible constant $\xi_2 > 0$ such that (f) holds follows by applying the next lemma with $\zeta_1 = \xi_1$.

LEMMA 9.35. *Fix $\zeta_1 \in (0, \chi]$. Then, there is a 0-admissible constant $\zeta_2 > 0$ (which also depends on ζ_1) such that $B_X(x_0, \zeta_2) \subseteq B_{X_{j_0}}(x_0, \zeta_1)$.*

Proof. Let $\zeta_2 \in (0, \zeta_1]$, we will pick ζ_2 at the end of the proof. Suppose $y \in B_X(x_0, \zeta_2)$, so that $\exists \gamma : [0, 1] \rightarrow B_X(x_0, \zeta_2)$ with $\gamma(0) = x_0$, $\gamma(1) = y$, $\gamma'(t) = \sum_{j=1}^q a_j(t) \zeta_2 X_j(\gamma(t))$, $\|\sum |a_j(t)|^2\|_{L^\infty([0,1])} < 1$. Let

$$t_0 = \sup\{t \in [0, 1] : \gamma(t') \in B_{X_{j_0}}(x_0, \zeta_1/2), \forall t' \in [0, t]\}.$$

We wish to show that if $\zeta_2 = \zeta_2(\zeta_1) > 0$ is taken to be a sufficiently small 0-admissible constant, then we have $t_0 = 1$ and $y = \gamma(1) \in B_{X_{j_0}}(x_0, \zeta_1)$.

In fact, we will prove $\gamma(t_0) \in B_{X_{j_0}}(x_0, \zeta_1/2)$. The result will then follow as if $t_0 < 1$, the fact that $B_{X_{j_0}}(x_0, \zeta_1/2)$ is open (see (c)) and γ is continuous show that $\gamma(t') \in B_{X_{j_0}}(x_0, \zeta_1/2)$ for $t' \in [0, t_0 + \epsilon)$ for some $\epsilon > 0$, which contradicts the choice of t_0 .

We turn to proving $\gamma(t_0) \in B_{X_{j_0}}(x_0, \zeta_1/2)$. We have

$$\begin{aligned} \gamma'(t) &= \sum_{k=1}^q a_k(t) \zeta_2 X_k(\gamma(t)) = \sum_{l=1}^n \left(\sum_{k=1}^q a_k(t) \zeta_2 \tilde{b}_k^l(\gamma(t)) \right) X_l(\gamma(t)) \\ &=: \sum_{l=1}^n \tilde{a}_l(t) \frac{\zeta_1}{2} X_l(\gamma(t)), \end{aligned}$$

where $\|\sum |\tilde{a}_l(t)|^2\|_{L^\infty([0, t_0])} \lesssim_0 \frac{\zeta_2}{\zeta_1}$ (see Lemma 9.33). Thus, by taking $\zeta_2 = \zeta_2(\zeta_1) > 0$ to be a sufficiently small 0-admissible constant, we have $\|\sum |\tilde{a}_l(t)|^2\|_{L^\infty([0, t_0])} < 1$. It follows that $\gamma(t_0) \in B_{X_{j_0}}(x_0, \zeta_1/2)$, which completes the proof. $\square$

Proof of (i). For $n+1 \leq k \leq q$, $1 \leq l \leq n$, set $b_k^l := \tilde{b}_k^l \circ \Phi$. Pulling back (9.53) via Φ shows $Y_k = \sum_{l=1}^n b_k^l Y_l$. The regularity of b_k^l now follows by combining (l) and the bounds in Lemma 9.33. $\square$

Proof of (h) for $n+1 \leq j \leq q$. This follows by combining (h) for $1 \leq j \leq n$ and (i). $\square$

Proof of (k). We prove the result for Zygmund spaces; the proof for Hölder spaces is similar, and we leave it to the reader. Let $s > 2$. The case $n = q$ of Theorem 4.7 gives $\|f\|_{\mathcal{C}^s(B^n(\eta_1))} \approx_{\{s-1, s-2\}} \|f\|_{\mathcal{C}_{Y_{J_0}}^s(B^n(\eta_1))}$. Also, $\|f\|_{\mathcal{C}_Y^s(B^n(\eta_1))} \approx_{\{s-1, s-2\}} \|f\|_{\mathcal{C}^s(B^n(\eta_1))}$ follows from Proposition 8.12, (g), (h), and the fact that η_1 is a $\{s-1, s-2\}$ -admissible constant, for $s > 2$. Here we are using $\nabla = (I + A)^{-1}Y_{J_0}$ and $\|(I + A)^{-1}\|_{\mathcal{C}^s(B^n(\eta_1); \mathbb{M}^{n \times n})} \lesssim_{\{s, s-1\}} 1$, for $s > 0$ (which follows from (g) and (h)). $\square$

9.4 Densities. In this section, we prove the results from Section 6. We recall the density ν_0 from (6.2), defined on $B_{X_{J_0}}(x_0, \chi)$:

$$\nu_0(x)(Z_1(x), \dots, Z_n(x)) := \left| \frac{Z_1(x) \wedge Z_2(x) \wedge \dots \wedge Z_n(x)}{X_1(x) \wedge X_2(x) \wedge \dots \wedge X_n(x)} \right|.$$

LEMMA 9.36. $\nu_0(X_1, \dots, X_n) \equiv 1$, and for $j_1, \dots, j_n \in \{1, \dots, q\}$, $\nu_0(X_{j_1}, \dots, X_{j_n}) \lesssim_0 1$.

Proof. That $\nu_0(X_1, \dots, X_n) \equiv 1$ follows directly from the definition. That $\nu_0(X_{j_1}, \dots, X_{j_n}) \lesssim_0 1$ follows from Theorem 4.7 (b). $\square$

LEMMA 9.37. Let V and W be n -dimensional real vector spaces, and let $A : W \rightarrow V$ be an invertible linear transformation. Let $v_1, \dots, v_n$ be a basis for V and let $w_1, \dots, w_n \in W$. Then,

$$\frac{Aw_1 \wedge Aw_2 \wedge \dots \wedge Aw_n}{v_1 \wedge v_2 \wedge \dots \wedge v_n} = \frac{w_1 \wedge w_2 \wedge \dots \wedge w_n}{A^{-1}v_1 \wedge A^{-1}v_2 \wedge \dots \wedge A^{-1}v_n}.$$

Proof. Let Z_1, Z_2 be one dimensional real vector spaces and let $B : Z_1 \rightarrow Z_2$ be an invertible linear transformation. Let $z_1 \in Z_1$ and $0 \neq z_2 \in Z_2$. We claim

$$\frac{Bz_1}{z_2} = \frac{z_1}{B^{-1}z_2}. \quad (9.54)$$

Indeed, let $\lambda_2 : Z_2 \rightarrow \mathbb{R}$ be any nonzero linear functional, and set $\lambda_1 := \lambda_2 \circ B : Z_1 \rightarrow \mathbb{R}$ so that λ_1 is also a nonzero linear functional. We have

$$\frac{Bz_1}{z_2} = \frac{\lambda_2(Bz_1)}{\lambda_2(z_2)} = \frac{\lambda_1(z_1)}{\lambda_1(B^{-1}z_2)} = \frac{z_1}{B^{-1}z_2}.$$

Applying (9.54) in the case $Z_1 = \bigwedge^n W$, $Z_2 = \bigwedge^n V$, and $B : Z_1 \rightarrow Z_2$ given by $B(w_1 \wedge w_2 \wedge \dots \wedge w_n) = (Aw_1) \wedge (Aw_2) \wedge \dots \wedge (Aw_n)$ completes the proof. $\square$

LEMMA 9.38. For $1 \leq j \leq n$, $\mathcal{L}_{X_j}\nu_0 = f_j^0\nu_0$, where $f_j^0 \in C(B_{X_{J_0}}(x_0, \chi))$. Furthermore, for $m \in \mathbb{N}$, $s \in [0, 1]$,

$$\|f_j^0\|_{C_{X_{J_0}}^{m,s}(B_{X_{J_0}}(x_0, \chi))} \lesssim_{\langle m, m, s \rangle} 1, \quad (9.55)$$

and for $s \in (0, \infty)$,

$$\|f_j^0\|_{\mathcal{C}_{X_{J_0}}^s(B_{X_{J_0}}(x_0, \chi))} \lesssim_{\{s, s\}} 1. \quad (9.56)$$

Proof. Set $\phi_t(x) = e^{tX_j}x$ so that $\mathcal{L}_{X_j}\nu_0 = \frac{\partial}{\partial t}\big|_{t=0}\phi_t^*\nu_0$. We write $d\phi_t(x)$ to denote the differential of ϕ_t in the x variable. We have, using Lemma 9.37,

$$\begin{aligned}
(\phi_t^*\nu_0)(x)(Z_1, \dots, Z_n) &= \nu_0(\phi_t(x))(d\phi_t(x)Z_1(x), \dots, d\phi_t(x)Z_n(x)) \\
&= \left| \frac{d\phi_t(x)Z_1(x) \wedge d\phi_t(x)Z_2(x) \wedge \dots \wedge d\phi_t(x)Z_n(x)}{X_1(\phi_t(x)) \wedge X_2(\phi_t(x)) \wedge \dots \wedge X_n(\phi_t(x))} \right| \\
&= \left| \frac{Z_1(x) \wedge Z_2(x) \wedge \dots \wedge Z_n(x)}{d\phi_t(x)^{-1}X_1(\phi_t(x)) \wedge d\phi_t(x)^{-1}X_2(\phi_t(x)) \wedge \dots \wedge d\phi_t(x)^{-1}X_n(\phi_t(x))} \right| \\
&= \left| \frac{Z_1(x) \wedge Z_2(x) \wedge \dots \wedge Z_n(x)}{\phi_t^*X_1(x) \wedge \phi_t^*X_2(x) \wedge \dots \wedge \phi_t^*X_n(x)} \right| \tag{9.57}
\end{aligned}$$

Fix $x \in B_{X_{j_0}}(x_0, \chi)$. We claim that the sign of

$$\frac{Z_1(x) \wedge Z_2(x) \wedge \dots \wedge Z_n(x)}{\phi_t^*X_1(x) \wedge \phi_t^*X_2(x) \wedge \dots \wedge \phi_t^*X_n(x)}$$

does not change for t small. To this end, let θ be a C^1 n -form which is nonzero near x . Since $X_1(x) \wedge X_2(x) \wedge \dots \wedge X_n(x) \neq 0$ (Theorem 4.7 (a)), $\theta(x)(X_1(x) \wedge X_2(x) \wedge \dots \wedge X_n(x)) \neq 0$, and so by continuity, for t small, $\theta(x)(\phi_t^*X_1(x) \wedge \phi_t^*X_2(x) \wedge \dots \wedge \phi_t^*X_n(x)) \neq 0$. We conclude that for t sufficiently small,

$$\frac{Z_1(x) \wedge Z_2(x) \wedge \dots \wedge Z_n(x)}{\phi_t^*X_1(x) \wedge \phi_t^*X_2(x) \wedge \dots \wedge \phi_t^*X_n(x)} = \frac{\theta(x)(Z_1(x) \wedge Z_2(x) \wedge \dots \wedge Z_n(x))}{\theta(x)(\phi_t^*X_1(x) \wedge \phi_t^*X_2(x) \wedge \dots \wedge \phi_t^*X_n(x))}$$

does not change sign, and is either never zero or always zero for small t .

Set, for t small,

$$\epsilon := \operatorname{sgn} \frac{Z_1(x) \wedge Z_2(x) \wedge \dots \wedge Z_n(x)}{\phi_t^*X_1(x) \wedge \phi_t^*X_2(x) \wedge \dots \wedge \phi_t^*X_n(x)},$$

and in the case the quantity inside sgn equals zero, the choice of ϵ does not matter. By the above discussion, ϵ does not depend on t (for t small). We have, using the functions $g_{j,J}^K$ from Lemmas 9.27 and 9.28,

$$\begin{aligned}
\frac{\partial}{\partial t}\bigg|_{t=0} (\phi_t^*\nu_0)(x)(Z_1(x), \dots, Z_n(x)) &= \frac{\partial}{\partial t}\bigg|_{t=0} \left| \frac{Z_1(x) \wedge Z_2(x) \wedge \dots \wedge Z_n(x)}{\phi_t^*X_1(x) \wedge \phi_t^*X_2(x) \wedge \dots \wedge \phi_t^*X_n(x)} \right| \\
&= \frac{\partial}{\partial t}\bigg|_{t=0} \epsilon \frac{Z_1(x) \wedge Z_2(x) \wedge \dots \wedge Z_n(x)}{\phi_t^*X_1(x) \wedge \phi_t^*X_2(x) \wedge \dots \wedge \phi_t^*X_n(x)} \\
&= \frac{\partial}{\partial t}\bigg|_{t=0} \epsilon \frac{\theta(x)(Z_1(x) \wedge Z_2(x) \wedge \dots \wedge Z_n(x))}{\theta(x)(\phi_t^*X_1(x) \wedge \phi_t^*X_2(x) \wedge \dots \wedge \phi_t^*X_n(x))} \\
&= -\epsilon \frac{\theta(x)(Z_1(x) \wedge Z_2(x) \wedge \dots \wedge Z_n(x))}{\theta(x)(\phi_t^*X_1(x) \wedge \phi_t^*X_2(x) \wedge \dots \wedge \phi_t^*X_n(x))^2} \frac{\partial}{\partial t} \theta(x)(\phi_t^*X_1(x) \wedge \phi_t^*X_2(x) \wedge \dots \wedge \phi_t^*X_n(x)) \bigg|_{t=0} \\
&= -\epsilon \frac{\theta(x)(Z_1(x) \wedge Z_2(x) \wedge \dots \wedge Z_n(x))}{\theta(x)(X_1(x) \wedge X_2(x) \wedge \dots \wedge X_n(x))} \frac{\theta(x)(\mathcal{L}_{X_j}(X_1 \wedge X_2 \wedge \dots \wedge X_n)(x))}{\theta(x)(X_1(x) \wedge X_2(x) \wedge \dots \wedge X_n(x))} \\
&= -\epsilon \frac{Z_1(x) \wedge Z_2(x) \wedge \dots \wedge Z_n(x)}{X_1(x) \wedge X_2(x) \wedge \dots \wedge X_n(x)} \frac{\mathcal{L}_{X_j}(X_1 \wedge X_2 \wedge \dots \wedge X_n)(x)}{X_1(x) \wedge X_2(x) \wedge \dots \wedge X_n(x)}
\end{aligned}$$

$$\begin{aligned}
&= -\frac{\mathcal{L}_{X_j}(X_1 \wedge X_2 \wedge \cdots \wedge X_n)(x)}{X_1(x) \wedge X_2(x) \wedge \cdots \wedge X_n(x)} \nu_0(x)(Z_1(x), \dots, Z_n(x)) \\
&= -\sum_{K \in \mathcal{I}_0(n, q)} g_{j, J_0}^K(x) \frac{\bigwedge X_K(x)}{\bigwedge X_{J_0}(x)} \nu_0(x)(Z_1(x), \dots, Z_n(x)).
\end{aligned}$$

We conclude that

$$f_j^0 = -\sum_{K \in \mathcal{I}_0(n, q)} g_{j, J_0}^K \frac{\bigwedge X_K}{\bigwedge X_{J_0}}.$$

(9.55) and (9.56) follow from Lemmas 9.27 and 9.32 and Proposition 8.3. $\square$

Let $\sigma_0 := \Phi^* \nu_0$, so that σ_0 is a density on $B^n(\eta_1)$. Define h_0 by $\sigma_0 = h_0 \sigma_{\text{Leb}}$, so that $h_0 \in C(B^n(\eta_1))$.

LEMMA 9.39. $h_0(t) = \det(I + A(t))^{-1}$, where A is the matrix from Theorem 4.7. In particular, $h_0(t) \approx_0 1$, $\forall t \in B^n(\eta_1)$. For $m \in \mathbb{N}$, $s \in [0, 1]$,

$$\|h_0\|_{C^{m, s}(B^n(\eta_1))} \lesssim_{\langle m, m-1, s \rangle} 1, \quad (9.58)$$

and for $s \in (0, \infty)$,

$$\|h_0\|_{\mathcal{C}^s(B^n(\eta_1))} \lesssim_{\{s, s-1\}} 1. \quad (9.59)$$

Proof. Because $\|A(t)\|_{\mathbb{M}^{n \times n}} \leq \frac{1}{2}$, $\forall t \in B^n(\eta_1)$ (Theorem 4.7 (g)), we have $|\det(I + A(t))^{-1}| = \det(I + A(t))^{-1}$, $\forall t \in B^n(\eta_1)$. We have,

$$\begin{aligned}
h_0(t) &= \sigma_0(t) \left(\frac{\partial}{\partial t_1}, \frac{\partial}{\partial t_2}, \dots, \frac{\partial}{\partial t_n} \right) = \sigma_0(t) ((I + A(t))^{-1} Y_1(t), \dots, (I + A(t))^{-1} Y_n(t)) \\
&= |\det(I + A(t))^{-1}| \sigma_0(t) (Y_1(t), \dots, Y_n(t)) \\
&= \det(I + A(t))^{-1} \nu_0(\Phi(t)) (X_1(\Phi(t)), \dots, X_n(\Phi(t))) \\
&= \det(I + A(t))^{-1}.
\end{aligned}$$

That $h_0(t) \approx_0 1$, $\forall t \in B^n(\eta_1)$ follows from the fact that $\|A(t)\|_{\mathbb{M}^{n \times n}} \leq \frac{1}{2}$, $\forall t \in B^n(\eta_1)$ (Theorem 4.7 (g)). Using Proposition 8.3 (applied to the cofactor representation of $(I + A)^{-1}$), (9.58) and (9.59) follow from the corresponding regularity for A as described in Theorem 4.7 (h)—here we are using that the regularity for A and the regularity for $Y_1, \dots, Y_n$ are the same, by the definition of A . $\square$

We now turn to studying the density ν from Section 6; thus we use the functions f_j from (6.1). Because ν_0 is a nonzero density on $B_{X_{J_0}}(x_0, \chi)$, there is a unique $g \in C(B_{X_{J_0}}(x_0, \chi))$ such that $\nu = g\nu_0$.

LEMMA 9.40. For $1 \leq j \leq n$, $X_j g = (f_j - f_j^0)g$.

Proof. We have,

$$f_j g \nu_0 = f_j \nu = \mathcal{L}_{X_j} \nu = \mathcal{L}_{X_j} (g \nu_0) = (X_j g) \nu_0 + g \mathcal{L}_{X_j} \nu_0 = (X_j g) \nu_0 + g f_j^0 \nu_0.$$

The result follows. $\square$

LEMMA 9.41. *Theorem 6.5 (i) holds. Namely,*

$$g(x) \approx_{0;\nu} g(x_0) = \nu(x_0)(X_1(x_0), \dots, X_n(x_0)), \quad \forall x \in B_{X_{j_0}}(x_0, \chi).$$

Proof. Note $g(x_0) = g(x_0)\nu_0(x_0)(X_1(x_0), \dots, X_n(x_0)) = \nu(x_0)(X_1(x_0), \dots, X_n(x_0))$, by definition. So it suffices to show $g(x) \approx_{0;\nu} g(x_0)$ for $x \in B_{X_{j_0}}(x_0, \chi)$.

Let $\gamma : [0, 1] \rightarrow B_{X_{j_0}}(x_0, \chi)$ be such that $\gamma(0) = x_0$, $\gamma(1) = x$, $\gamma'(t) = \sum_{j=1}^n a_j(t)\chi X_j(\gamma(t))$, $\|\sum |a_j(t)|^2\|_{L^\infty([0,1])} < 1$. We have, using Lemma 9.40,

$$\frac{d}{dt}g(\gamma(t)) = \sum_{j=1}^n a_j(t)\chi(X_j g)(\gamma(t)) = \sum_{j=1}^n a_j(t)\chi(f_j(\gamma(t)) - f_j^0(\gamma(t)))g(\gamma(t)).$$

Hence, $g(\gamma(t))$ satisfies an ODE. Solving this ODE we have

$$g(x) = g(\gamma(1)) = e^{\int_0^1 \sum_{j=1}^n a_j(s)\chi(f_j(\gamma(s)) - f_j^0(\gamma(s))) ds} g(x_0).$$

We know $\|f_j^0\|_{C(B_{X_{j_0}}(x_0, \chi))} \lesssim_0 1$ (by the case $m = 0$, $s = 0$ of (9.55)). Using this and the definition of $0;\nu$ -admissible constants, $g(x) \approx_{0;\nu} g(x_0)$ follows immediately, completing the proof. $\square$

LEMMA 9.42. *Theorem 6.5 (ii) holds. Namely, for $m \in \mathbb{N}$, $s \in [0, 1]$,*

$$\|g\|_{C_{X_{j_0}}^{m,s}(B_{X_{j_0}}(x_0, \chi))} \lesssim_{\langle m-1, m-1, s; \nu \rangle} |\nu(X_1, \dots, X_n)(x_0)|, \quad (9.60)$$

and for $s \in (0, \infty)$,

$$\|g\|_{\mathcal{C}_{X_{j_0}}^s(B_{X_{j_0}}(x_0, \chi))} \lesssim_{\{s-1, s-1; \nu\}} |\nu(X_1, \dots, X_n)(x_0)|. \quad (9.61)$$

Proof. We begin with (9.60). First note that

$$\|g\|_{C(B_{X_{j_0}}(x_0, \chi))} \lesssim_{0;\nu} |\nu(X_1, \dots, X_n)(x_0)|, \quad (9.62)$$

which follows immediately from Lemma 9.41. We claim that

$$\|g\|_{C_{X_{j_0}}^1(B_{X_{j_0}}(x_0, \chi))} \lesssim_{0;\nu} |\nu(X_1, \dots, X_n)(x_0)|. \quad (9.63)$$

Indeed, using Lemma 9.40, for each $1 \leq j \leq n$,

$$\begin{aligned} \|X_j g\|_{C(B_{X_{j_0}}(x_0, \chi))} &= \|(f_j - f_j^0)g\|_{C(B_{X_{j_0}}(x_0, \chi))} \\ &\lesssim_{0;\nu} \|g\|_{C(B_{X_{j_0}}(x_0, \chi))} \lesssim_{0;\nu} |\nu(X_1, \dots, X_n)(x_0)|, \end{aligned} \quad (9.64)$$

where in the last inequality we have used (9.62) and in the second to last inequality we have used $\|f_j\|_{C(B_{X_{j_0}}(x_0, \chi))} \lesssim_{0;\nu} 1$ (which follows from the definition of $0;\nu$ -admissible constants) and $\|f_j^0\|_{C(B_{X_{j_0}}(x_0, \chi))} \lesssim_0 1$ (which follows from the case $m = 0$, $s = 0$ of (9.55)). Combining (9.62) and (9.64) proves (9.63).

We prove (9.60) by induction on m . For the base case, $m = 0$, we have using Lemma 8.1 (i) and (ii), and (9.63),

$$\|g\|_{C_{X_{J_0}}^{0,s}(B_{X_{J_0}}(x_0,\chi))} \leq 3\|g\|_{C_{X_{J_0}}^{0,1}(B_{X_{J_0}}(x_0,\chi))} \leq 3\|g\|_{C_{X_{J_0}}^1(B_{X_{J_0}}(x_0,\chi))} \lesssim_{0;\nu} |\nu(X_1, \dots, X_n)(x_0)|.$$

This proves the case $m = 0$ of (9.60).

We now assume (9.60) for $m - 1$ and prove it for m . We have

$$\|g\|_{C_{X_{J_0}}^{m,s}(B_{X_{J_0}}(x_0,\chi))} = \|g\|_{C_{X_{J_0}}^{m-1,s}(B_{X_{J_0}}(x_0,\chi))} + \sum_{j=1}^n \|X_j g\|_{C_{X_{J_0}}^{m-1,s}(B_{X_{J_0}}(x_0,\chi))}.$$

The first term is $\lesssim_{\langle m-2, m-2, s; \nu \rangle} |\nu(X_1, \dots, X_n)(x_0)|$ by the inductive hypothesis, so we focus only on the second term. We have, using Lemma 9.40 and Proposition 8.3, for a constant C_m depending only on m , for $1 \leq j \leq n$,

$$\begin{aligned} \|X_j g\|_{C_{X_{J_0}}^{m-1,s}(B_{X_{J_0}}(x_0,\chi))} &= \|(f_j - f_j^0)g\|_{C_{X_{J_0}}^{m-1,s}(B_{X_{J_0}}(x_0,\chi))} \\ &\leq C_m \|f_j - f_j^0\|_{C_{X_{J_0}}^{m-1,s}(B_{X_{J_0}}(x_0,\chi))} \|g\|_{C_{X_{J_0}}^{m-1,s}(B_{X_{J_0}}(x_0,\chi))} \\ &\lesssim_{\langle m-1, m-1, s; \nu \rangle} |\nu(X_1, \dots, X_n)(x_0)|, \end{aligned}$$

where the last inequality follows from the inductive hypothesis, (9.55), and the definition of $\langle m - 1, m - 1, s; \nu \rangle$ -admissible constants. (9.60) follows. $\square$

We turn to (9.61), which we prove by induction on m , where $s \in (m, m + 1]$. We begin with the base case, $m = 0$, so that $s \in (0, 1]$. Using Lemma 8.1 (iii) and (9.60) we have

$$\|g\|_{\mathcal{C}_{X_{J_0}}^s(B_{X_{J_0}}(x_0,\chi))} \leq 5\|g\|_{C_{X_{J_0}}^{0,s}(B_{X_{J_0}}(x_0,\chi))} \lesssim_{0;\nu} |\nu(X_1, \dots, X_n)(x_0)|.$$

(9.61) follows for $s \in (0, 1]$. From here the inductive step follows just as in the inductive step for (9.60), and we leave the details to the reader. $\square$

LEMMA 9.43. *Let $h(t)$ be as in Theorem 6.5. Then $h(t) = h_0(t)g \circ \Phi(t)$.*

Proof. We have

$$\Phi^* \nu = \Phi^* g \nu_0 = (g \circ \Phi) \Phi^* \nu_0 = (g \circ \Phi) h_0 \sigma_{\text{Leb}},$$

completing the proof. $\square$

Proof of Theorem 6.5 (iii). This follows from Lemmas 9.39, 9.41 and 9.43. $\square$

Proof of Theorem 6.5 (iv). We prove the result for Zygmund spaces; the same proof works for Hölder spaces, and we leave the details to the reader. Using Theorem 4.7 (1) we have

$$\|g \circ \Phi\|_{\mathcal{C}^s(B^n(\eta_1))} \lesssim_{\{s-1, s-2\}} \|g\|_{\mathcal{C}_{X_{J_0}}^s(B_{X_{J_0}}(x_0,\chi))} \lesssim_{\{s-1, s-1; \nu\}} |\nu(X_1, \dots, X_n)(x_0)|, \quad (9.65)$$

where the last inequality uses (9.61). Since $h(t) = h_0(t)g \circ \Phi(t)$ (Lemma 9.43), combining (9.65) and (9.59), and using Proposition 8.3 completes the proof. $\square$

Having completed the proof of Theorem 6.5, we turn to Corollary 6.6. To facilitate this, we introduce a corollary of Theorem 4.7.

COROLLARY 9.44. *Let η_1, ξ_1, ξ_2 be as in Theorem 4.7. Then, there exist 1-admissible constants $0 < \eta_2 \leq \eta_1$, $0 < \xi_4 \leq \xi_3 \leq \xi_2$ such that*

$$\begin{aligned} B_X(x_0, \xi_4) &\subseteq B_{X_{j_0}}(x_0, \xi_3) \subseteq \Phi(B^n(\eta_2)) \subseteq B_{X_{j_0}}(x_0, \xi_2) \subseteq B_X(x_0, \xi_2) \\ &\subseteq B_{X_{j_0}}(x_0, \xi_1) \subseteq \Phi(B^n(\eta_1)) \subseteq B_{X_{j_0}}(x_0, \chi) \subseteq B_X(x_0, \xi). \end{aligned}$$

Proof. After obtaining η_1, ξ_1, ξ_2 from Theorem 4.7, apply Theorem 4.7 again with ξ replaced by ξ_2 to obtain η_2, ξ_3 , and ξ_4 as in the statement of the corollary. $\square$

Proof of Corollary 6.6. We have

$$\begin{aligned} \nu(B_{X_{j_0}}(x_0, \xi_2)) &= \int_{B_{X_{j_0}}(x_0, \xi_2)} \nu = \int_{\Phi^{-1}(B_{X_{j_0}}(x_0, \xi_2))} \Phi^* \nu \\ &= \int_{\Phi^{-1}(B_{X_{j_0}}(x_0, \xi_2))} h(t) dt \approx_{0; \nu} \text{Vol}(\Phi^{-1}(B_{X_{j_0}}(x_0, \xi_2))) \nu(X_1, \dots, X_n)(x_0), \end{aligned} \quad (9.66)$$

where $\text{Vol}(\cdot)$ denotes Lebesgue measure, and we have used Theorem 6.5 (iii). By Corollary 9.44, and the fact that $\eta_1, \eta_2 > 0$ are 1-admissible constants, we have

$$1 \approx_1 \text{Vol}(B^n(\eta_2)) \leq \text{Vol}(\Phi^{-1}(B_{X_{j_0}}(x_0, \xi_2))) \leq \text{Vol}(B^n(\eta_1)) \approx_1 1. \quad (9.67)$$

Combining (9.66) and (9.67) proves $\nu(B_{X_{j_0}}(x_0, \xi_2)) \approx_{1; \nu} \nu(X_1, \dots, X_n)(x_0)$. The same proof works with $B_{X_{j_0}}(x_0, \xi_2)$ replaced by $B_X(x_0, \xi_2)$, which completes the proof of (6.3).

All that remains to prove (6.4) is to show

$$|\nu(X_1, \dots, X_n)(x_0)| \approx_0 \max_{(j_1, \dots, j_n) \in \mathcal{I}(n, q)} |\nu(X_{j_1}, \dots, X_{j_n})(x_0)|.$$

We have, using Lemma 9.36,

$$\begin{aligned} |\nu(X_1, \dots, X_n)(x_0)| &= |g(x_0) \nu_0(X_1, \dots, X_n)(x_0)| \\ &= |g(x_0)| \approx_0 |g(x_0)| \max_{(j_1, \dots, j_n) \in \mathcal{I}(n, q)} |\nu_0(X_{j_1}, \dots, X_{j_n})(x_0)| \\ &= \max_{(j_1, \dots, j_n) \in \mathcal{I}(n, q)} |g(x_0) \nu_0(X_{j_1}, \dots, X_{j_n})(x_0)| \\ &= \max_{(j_1, \dots, j_n) \in \mathcal{I}(n, q)} |\nu(X_{j_1}, \dots, X_{j_n})(x_0)|, \end{aligned}$$

completing the proof. $\square$

9.5 More on the assumptions. In this section we prove Proposition 4.14. The existence of $\eta > 0$ as in Proposition 4.14 follows immediately from the Picard–Lindelöf Theorem, so we focus on the existence of $\delta_0 > 0$. The key is the next lemma.

LEMMA 9.45. *Suppose Z is a C^1 vector field on an open set $V \subseteq \mathbb{R}^n$. Then, there exists $\delta > 0$, depending only on n , such that if $\|Z\|_{C^1(V; \mathbb{R}^n)} \leq \delta$, then there does not exist $x \in V$ with:*

- $e^{tZ}x \in V$, $\forall t \in [0, 1]$.
- $e^Z x = x$.
- $Z(x) \neq 0$.

Proof. For a proof of this classical result, see [Str11, Lemma 3.19]. □

To prove the existence of δ_0 as in Proposition 4.14, since K is compact, it suffices to prove the next lemma.

LEMMA 9.46. *Let $X_1, \dots, X_q$ be C^1 vector fields on a C^2 manifold $\mathfrak{M}$. For all $x \in \mathfrak{M}$, there exists an open set $N \subseteq \mathfrak{M}$ with $x \in N$, and $\delta_0 > 0$ such that $\forall \theta \in S^{q-1}$ if $y \in N$ is such that $\theta_1 X_q(y) + \dots + \theta_q X_q(y) \neq 0$, then $\forall r \in (0, \delta_0]$,*

$$e^{r\theta_1 X_1 + \dots + r\theta_q X_q} y \neq y.$$

Proof. Since this result is local, it suffices to prove the lemma in the case when $\mathfrak{M} = B^n(1)$ and $x = 0 \in \mathbb{R}^n$. We set $N := B^n(1/2)$. Take $\delta = \delta(n) > 0$ as in Lemma 9.45. Take $\delta_1 > 0$ so small that $\forall y \in B^n(1/2)$, $t \in B^q(\delta_1)$, we have $e^{t_1 X_1 + \dots + t_q X_q} y \in B^n(3/4)$. Set $C := \max_{1 \leq j \leq q} \|X_j\|_{C^1(B^n(3/4); \mathbb{R}^n)}$, and let $\delta_0 = \min\{\delta_1, \delta/qC\}$. From here, the result follows from Lemma 9.45. □

Acknowledgments

We thank the referee whose detailed comments improved the exposition. Stovall was partially supported by National Science Foundation Grant No. 1600458. Street was partially supported by National Science Foundation Grant Nos. 1401671 and 1764265. This material is partially based upon work supported by the National Science Foundation under Grant No. 1440140, while the authors were in residence at the Mathematical Sciences Research Institute in Berkeley, California, during the spring semester of 2017.

A Proof of Proposition 3.1

The ideas behind Proposition 3.1 are well-known to experts; however, we could not find an exact statement of Proposition 3.1 in the literature, so we include the proof here for completeness, with the understanding that the methods used are known to experts. It seems

closely related to the theory of orbits of Sussman [Sus73] and Stefan [Ste74], though does not follow directly from these theories. Similar methods have been used to prove the Frobenius theorem for Lipschitz vector fields; see [MM13a] and references therein.

We begin with the existence of the C^2 structure; we take all the same notation as in the statement of Proposition 3.1. Set $D := \dim \mathfrak{M}$, and let $(\phi_\alpha, U_\alpha)_{\alpha \in \mathcal{A}}$ be a C^2 atlas for $\mathfrak{M}$ with $\{U_\alpha : \alpha \in \mathcal{A}\}$ an open cover for $\mathfrak{M}$ and $\phi_\alpha : U_\alpha \rightarrow B^D(1)$ a C^2 diffeomorphism.

Let $X_j^{(\alpha)} = (\phi_\alpha)_* X_j$ so that $X_j^{(\alpha)}$ is a C^1 vector field on $B^D(1)$. We may pick the above atlas so that $\|X_j^{(\alpha)}\|_{C^1(B^D(1); \mathbb{R}^n)} < \infty$.

LEMMA A.1. *Let Z be as in the beginning of Section 3.1. The topology on Z (induced by the metric ρ) is finer than the topology as a subspace of $\mathfrak{M}$.*

Proof. Let $U \subseteq \mathfrak{M}$ be an open set and let $x \in U \cap Z$. We wish to show that there is a $\delta > 0$ with $B_X(x, \delta) \subseteq U$. Since $x \in U_\alpha$ for some $\alpha \in \mathcal{A}$, we may replace U with $U \cap U_\alpha$, and therefore assume $U \subseteq U_\alpha$ for some $\alpha \in \mathcal{A}$.

By the Picard-Lindelöf Theorem, there exists $\delta > 0$ so small such that given $a_1, \dots, a_q \in L^\infty([0, 1])$ with $\|\sum |a_j|^2\|_{L^\infty([0, 1])} < 1$, there exists a unique $\tilde{\gamma} : [0, 1] \rightarrow \phi_\alpha(U)$ with

$$\tilde{\gamma}(0) = \phi_\alpha(x) \text{ and } \tilde{\gamma}'(t) = \sum_{j=1}^q a_j(t) \delta X^{(\alpha)}(\tilde{\gamma}(t)). \quad (\text{A.1})$$

We claim $B_X(x, \delta) \subseteq U$. Indeed, fix $y \in B_X(x, \delta)$. By the definition of $B_X(x, \delta)$, $\exists \gamma : [0, 1] \rightarrow B_X(x, \delta)$, $\gamma(0) = x$, $\gamma(1) = y$, $\gamma'(t) = \sum_{j=1}^q a_j(t) \delta X_j(\gamma(t))$. Let $\tilde{\gamma} : [0, 1] \rightarrow \phi_\alpha(U)$ be the unique solution to (A.1) with this choice of $a_1, \dots, a_q$, and set $\hat{\gamma} := \phi_\alpha^{-1} \circ \tilde{\gamma}$. Then, $\hat{\gamma}(0) = x = \gamma(0)$, $\hat{\gamma}'(t) = \sum_{j=1}^q a_j(t) \delta X_j(\gamma(t)) = \gamma'(t)$. Standard uniqueness theorems for ODEs show $\gamma = \hat{\gamma}$, and therefore $y = \gamma(1) = \hat{\gamma}(1) = \phi_\alpha^{-1}(\tilde{\gamma}(1))$. Since $\tilde{\gamma}(1) \in \phi_\alpha(U)$, it follows $y \in U$, which completes the proof. $\square$

Recall, M is a connected open subset of Z which is given the topology as subspace of Z ; i.e., M is given the topology induced by the metric ρ .

Set $M_\alpha := \phi_\alpha(U_\alpha \cap M)$; we give M_α the topology so that $\phi_\alpha : M \cap U_\alpha \rightarrow M_\alpha$ is a homeomorphism (with $M \cap U_\alpha \subseteq M$ given the topology as a subspace of M). Let $X^{(\alpha)}(u)$ denote the $D \times q$ matrix $X^{(\alpha)}(u) = (X_1^{(\alpha)}(u) | \dots | X_q^{(\alpha)}(u))$. For $K = (k_1, \dots, k_l) \in \mathcal{I}(l, q)$ let $X_K^{(\alpha)}$ denote the list of vector fields $X_{k_1}^{(\alpha)}, \dots, X_{k_l}^{(\alpha)}$ and for $J = (j_1, \dots, j_l) \in \mathcal{I}(l, D)$ let $X_{J,K}^{(\alpha)}$ denote the $l \times l$ submatrix of $X^{(\alpha)}(u)$ given by taking the rows listed in J and the columns listed in K .

LEMMA A.2. *For $u \in M_\alpha$, $1 \leq k \leq q$, $1 \leq l \leq \min\{q, D\}$, $K \in \mathcal{I}(l, q)$, $J \in \mathcal{I}(l, D)$*

$$X_k^{(\alpha)} \det X_{J,K}^{(\alpha)}(u) = \sum_{\substack{K' \in \mathcal{I}(l, q) \\ J' \in \mathcal{I}(l, D)}} f_{k,J,K}^{J',K'} \det X_{J',K'}^{(\alpha)}(u),$$

where $f_{k,J,K}^{J',K'} : M_\alpha \rightarrow \mathbb{R}$ are locally bounded.

Proof. Let $J = (j_1, \dots, j_l)$, $K = (k_1, \dots, k_l)$. Then, $\det X_{J,K}^{(\alpha)} = \nu_J(X_{k_1}^{(\alpha)}, \dots, X_{k_l}^{(\alpha)})$, where ν_J is the l -form $du_{j_1} \wedge du_{j_2} \wedge \dots \wedge du_{j_l}$. Hence, using [Lee03, Proposition 18.9] we have

$$\begin{aligned}
X_k^{(\alpha)} \det X_{J,K}^{(\alpha)} &= \mathcal{L}_{X_k^{(\alpha)}} \left(\nu_J(X_{k_1}^{(\alpha)}, \dots, X_{k_l}^{(\alpha)}) \right) \\
&= \left(\mathcal{L}_{X_k^{(\alpha)}} \nu_J \right) (X_{k_1}^{(\alpha)}, \dots, X_{k_l}^{(\alpha)}) + \nu_J([X_k^{(\alpha)}, X_{k_1}^{(\alpha)}], X_{k_2}^{(\alpha)}, \dots, X_{k_l}^{(\alpha)}) \\
&\quad + \nu_J(X_{k_1}^{(\alpha)}, [X_k^{(\alpha)}, X_{k_2}^{(\alpha)}], X_{k_3}^{(\alpha)}, \dots, X_{k_l}^{(\alpha)}) + \dots + \nu_J(X_{k_1}^{(\alpha)}, \dots, X_{k_{l-1}}^{(\alpha)}, [X_k^{(\alpha)}, X_{k_l}^{(\alpha)}])
\end{aligned} \tag{A.2}$$

We begin with the first term on the right hand side of (A.2). Since $X_k^{(\alpha)}$ is a C^1 vector field, $\mathcal{L}_{X_k^{(\alpha)}} \nu_J$ is a C^0 l -form on $B^D(1)$ and we have

$$\mathcal{L}_{X_k^{(\alpha)}} \nu_J = \sum_{J' \in \mathcal{I}(l, D)} f_{k,J}^{J'} \nu_{J'},$$

where $\|f_{k,J}^{J'}\|_{C^0(B^D(1))} < \infty$. Hence

$$\left(\mathcal{L}_{X_k^{(\alpha)}} \nu_J \right) (X_{k_1}^{(\alpha)}, \dots, X_{k_l}^{(\alpha)}) = \sum_{J' \in \mathcal{I}(l, D)} f_{k,J}^{J'} \det X_{J',K}^{(\alpha)},$$

as desired.

We now turn to the rest of the terms on the right hand side of (A.2). These terms are all similar, so we only discuss the first. We have

$$\begin{aligned}
\nu_J([X_k^{(\alpha)}, X_{k_1}^{(\alpha)}], X_{k_2}^{(\alpha)}, \dots, X_{k_l}^{(\alpha)}) &= \sum_r (c_{m,k_1}^r \circ \phi_\alpha) \nu_J(X_r^{(\alpha)}, X_{k_2}^{(\alpha)}, \dots, X_{k_l}^{(\alpha)}) \\
&= \sum_r (c_{m,k_1}^r \circ \phi_\alpha) \det X_{J,K_r}^{(\alpha)},
\end{aligned}$$

where $K_r = (r, k_2, \dots, k_l) \in \mathcal{I}(l, q)$. The result follows. $\square$

For $1 \leq l \leq \min\{D, q\}$ let $\det_{l \times l} X^{(\alpha)}(u)$ denote the vector whose components are $\det X_{J,K}^{(\alpha)}(u)$, where $J \in \mathcal{I}(l, D)$, $K \in \mathcal{I}(l, q)$.

LEMMA A.3. For $u \in M_\alpha$, $1 \leq j \leq q$, $1 \leq l \leq \min\{D, q\}$, $J \in \mathcal{I}(l, D)$, $K \in \mathcal{I}(l, q)$,

$$\left| X_j^{(\alpha)} \det X_{J,K}^{(\alpha)}(u) \right| \leq g_{j,J,K}(u) \left| \det_{l \times l} X^{(\alpha)}(u) \right|$$

where $g_{j,J,K} : M_\alpha \rightarrow [0, \infty)$ is locally bounded.

Proof. This follows immediately from Lemma A.2. $\square$

LEMMA A.4. Let $\gamma : [0, 1] \rightarrow M_\alpha$ be such that $\gamma'(t) = \sum_{j=1}^q a_j(t) X_j^{(\alpha)}(\gamma(t))$, where $a_j \in L^\infty([0, 1])$. Then,

$$\dim \text{span}\{X_1^{(\alpha)}(\gamma(0)), \dots, X_q^{(\alpha)}(\gamma(0))\} = \dim \text{span}\{X_1^{(\alpha)}(\gamma(1)), \dots, X_q^{(\alpha)}(\gamma(1))\}.$$

Proof. We will show

$$\left| \det_{l \times l} X^{(\alpha)}(\gamma(0)) \right| = 0 \Rightarrow \left| \det_{l \times l} X^{(\alpha)}(\gamma(1)) \right| = 0. \tag{A.3}$$

To see why (A.3) implies the result note that by reversing γ , we have

$$\left| \det_{l \times l} X^{(\alpha)}(\gamma(0)) \right| = 0 \Leftrightarrow \left| \det_{l \times l} X^{(\alpha)}(\gamma(1)) \right| = 0,$$

and by noting that $\dim \text{span}\{X_1^{(\alpha)}(u), \dots, X_q^{(\alpha)}(u)\} \geq l \Leftrightarrow |\det_{l \times l} X^{(\alpha)}(u)| \neq 0$, the result follows. We turn to proving (A.3). We have, using Lemma A.3,

$$\begin{aligned} \frac{d}{dt} \left| \det_{l \times l} X^{(\alpha)}(\gamma(t)) \right|^2 &= 2 \sum_{\substack{J \in \mathcal{I}(l, D) \\ K \in \mathcal{I}(l, q)}} \det X_{J,K}^{(\alpha)}(\gamma(t)) \frac{d}{dt} \det X_{J,K}^{(\alpha)}(\gamma(t)) \\ &= 2 \sum_{\substack{J \in \mathcal{I}(l, D) \\ K \in \mathcal{I}(l, q)}} \det X_{J,K}^{(\alpha)}(\gamma(t)) \left(\sum_{j=1}^q a_j(t) \left(X_j^{(\alpha)} \det X_{J,K}^{(\alpha)} \right) (\gamma(t)) \right) \\ &\leq 2 \sum_{\substack{J \in \mathcal{I}(l, D) \\ K \in \mathcal{I}(l, q)}} \left(\sup_{t \in [0,1]} g_{j,J,K}(\gamma(t)) \right) \left(\sum_{j=1}^q \|a_j\|_{L^\infty([0,1])} \right) \left| \det_{l \times l} X^{(\alpha)}(\gamma(t)) \right|^2. \end{aligned}$$

We conclude,

$$\frac{d}{dt} \left| \det_{l \times l} X^{(\alpha)}(\gamma(t)) \right|^2 \leq C \left| \det_{l \times l} X^{(\alpha)}(\gamma(t)) \right|^2,$$

for some constant C . (A.3) follows by Grönwall's inequality. $\square$

PROPOSITION A.5. *The map $x \mapsto \dim \text{span}\{X_1(x), \dots, X_q(x)\}$, $M \rightarrow \mathbb{N}$ is constant.*

Proof. Since M is connected, it suffices to show the map is locally constant. Fix $x \in M$ and pick $\alpha \in \mathcal{A}$ such that $x \in U_\alpha$. Take $\delta > 0$ so small that $B_X(x, \delta) \subset M \cap U_\alpha$ (here, we are using Lemma A.1). We wish to show $x \mapsto \dim \text{span}\{X_1(x), \dots, X_q(x)\}$, $B_X(x, \delta) \rightarrow \mathbb{N}$ is constant.

Take $y \in B_X(x, \delta)$, so that $\exists \gamma : [0, 1] \rightarrow \mathfrak{M}$, $\gamma(0) = x$, $\gamma(1) = y$, $\gamma'(t) = \sum a_j(t) \delta X_j(\gamma(t))$, $\|\sum |a_j(t)|^2\|_{L^\infty([0,1])} < 1$. Note, $\forall t \in [0, 1]$, $\gamma(t) \in B_X(x, \delta) \subseteq U_\alpha$.

Set $\tilde{\gamma}(t) := \phi_\alpha \circ \gamma(t)$. $\tilde{\gamma}$ satisfies all the hypotheses of Lemma A.4 and this shows

$$\dim \text{span}\{X_1^{(\alpha)}(\phi_\alpha(x)), \dots, X_q^{(\alpha)}(\phi_\alpha(x))\} = \dim \text{span}\{X_1^{(\alpha)}(\phi_\alpha(y)), \dots, X_q^{(\alpha)}(\phi_\alpha(y))\}.$$

Hence, $\dim \text{span}\{X_1(x), \dots, X_q(x)\} = \dim \text{span}\{X_1(y), \dots, X_q(y)\}$, completing the proof. $\square$

Set $n := \dim \text{span}\{X_1(x), \dots, X_q(x)\}$, $x \in M$ (by Proposition A.5, n does not depend on x).

LEMMA A.6. *Let $x \in M$ and $K = (k_1, \dots, k_n) \in \mathcal{I}(n, q)$ be such that $X_{k_1}(x), \dots, X_{k_n}(x)$ are linearly independent. Then, there is an open set $U \subseteq \mathfrak{M}$, containing x , $J \in \mathcal{I}(n, D)$, and $\delta > 0$ such that the following hold:*

- (i) $B_X(x, \delta) \subseteq U$.
- (ii) $\exists \alpha \in \mathcal{A}$, $U \subseteq U_\alpha$.
- (iii) $\inf_{u \in \phi_\alpha(U)} \left| \det X_{J,K}^{(\alpha)}(u) \right| > 0$.

- (iv) $\forall y \in B_X(x, \delta)$, $\text{span}\{X_{k_1}(y), \dots, X_{k_n}(y)\} = \text{span}\{X_1(y), \dots, X_q(y)\}$.
- (v) $\forall y \in B_X(x, \delta)$, $[X_{k_i}, X_{k_j}](y) \in \text{span}\{X_{k_1}(y), \dots, X_{k_n}(y)\}$.
- (vi) For $1 \leq j \leq q$, $1 \leq l \leq n$, $\exists b_j^l \in C^1(U)$, $\|b_j^l \circ \phi_\alpha^{-1}\|_{C^1(\phi_\alpha(U))} < \infty$, such that $\forall y \in B_X(x, \delta)$,

$$X_j(y) = \sum_{l=1}^n b_j^l(y) X_{k_l}(y). \quad (\text{A.4})$$

Proof. Let $U \subseteq \mathfrak{M}$ be a neighborhood of x which may shrink from line to line. First, we may take U so small that $U \subseteq U_\alpha$ for some $\alpha \in \mathcal{A}$. Since $X_{k_1}^{(\alpha)}(\phi_\alpha(x)), \dots, X_{k_n}^{(\alpha)}(\phi_\alpha(x))$ are linearly independent, by the hypotheses, $\exists J \in \mathcal{I}(n, D)$ such that

$$\left| \det X_{J,K}^{(\alpha)}(\phi_\alpha(x)) \right| > 0.$$

By the continuity of the map $u \mapsto \left| \det X_{J,K}^{(\alpha)}(u) \right|$, we may shrink U so that (iii) holds. We take $\delta > 0$ so small that (i) holds; here we are using Lemma A.1.

Since $\forall u \in \phi_\alpha(U)$, $\left| \det X_{J,K}^{(\alpha)}(u) \right| > 0$ we have $\forall y \in B_X(x, \delta) \subseteq U$,

$$\dim \text{span}\{X_{k_1}(y), \dots, X_{k_n}(y)\} = n = \dim \text{span}\{X_1(y), \dots, X_q(y)\},$$

proving (iv).

Since $[X_{k_i}, X_{k_j}](y) \in \text{span}\{X_1(y), \dots, X_q(y)\}$, $\forall y \in M$ (by assumption), (v) follows from (iv).

Finally, for (vi), set

$$b_j^l(y) := \frac{\det X_{J, K_{j,l}}^{(\alpha)}(\phi_\alpha(y))}{\det X_{J,K}^{(\alpha)}(\phi_\alpha(y))},$$

where $K_{j,l}$ is the same as K but with k_l repalced by j . That $\|b_j^l \circ \phi_\alpha^{-1}\|_{C^1(\phi_\alpha(U))} < \infty$ follows from (iii) and the fact that $X_1, \dots, X_q \in C^1$. (A.4) follows from Cramer's rule. $\square$

PROPOSITION A.7. *Let $x \in M$. Then there exist an open set $U \subseteq \mathfrak{M}$, containing x , $\delta > 0$, and C^1 vector fields $V_1, \dots, V_n$ on U such that the following hold:*

- (i) $B_X(x, \delta) \subseteq U$.
- (ii) $\exists \alpha \in \mathcal{A}$, $U \subseteq U_\alpha$.
- (iii) For $1 \leq j \leq q$, $1 \leq l \leq n$, $\exists f_j^l \in C^1(U)$, $\|f_j^l \circ \phi_\alpha^{-1}\|_{C^1(\phi_\alpha(U))} < \infty$ such that $\forall y \in B_X(x, \delta)$,

$$X_j(y) = \sum_{l=1}^n f_j^l(y) V_l(y).$$

- (iv) $\forall y \in B_X(x, \delta)$, $V_1(y), \dots, V_n(y)$ are linearly independent.

- (v) For all $1 \leq l \leq n$, $1 \leq j \leq q$, $\exists g_l^j \in C^1(U)$, $\|g_l^j \circ \phi_\alpha^{-1}\|_{C^1(\phi_\alpha(U))} < \infty$, such that $\forall y \in B_X(x, \delta)$,

$$V_l(y) = \sum_{j=1}^q g_l^j(y) X_j(y).$$

- (vi) $\forall y \in B_X(x, \delta)$, $[V_j, V_k](y) = 0$, $1 \leq j, k \leq n$.

Proof. Take $K = (k_1, \dots, k_n) \in \mathcal{I}(n, q)$ such that $X_{k_1}(x), \dots, X_{k_n}(x)$ are linearly independent and let $J \in \mathcal{I}(n, D)$, $U \subseteq \mathfrak{M}$, $\delta > 0$ be as in Lemma A.6. Without loss of generality, we may reorder the vector fields and coordinates so that $J = (1, \dots, n)$, $K = (1, \dots, n)$. For $1 \leq j \leq n$, $u \in \phi_\alpha(U)$, write

$$X_j^{(\alpha)} = \sum_{k=1}^D h_{j,k} \circ \phi_\alpha^{-1}(u) \frac{\partial}{\partial x_k},$$

and let $H(y)$ denote the $n \times n$ matrix $H(y) = (h_{j,k})_{1 \leq j \leq n, 1 \leq k \leq n}$. Clearly, $\|h_{j,k} \circ \phi_\alpha^{-1}\|_{C^1(\phi_\alpha(U))} < \infty$.

By Lemma A.6 (iii), $\inf_{u \in \phi_\alpha(U)} |\det H(u)| > 0$. Define $h^{j,k}$ by $H(y)^{-1} = (h^{j,k}(y))_{1 \leq j \leq n, 1 \leq k \leq n}$, $y \in U$. By the above comments, $\|h^{j,k} \circ \phi_\alpha^{-1}\|_{C^1(\phi_\alpha(U))} < \infty$. Set

$$V_j(y) = \sum_{k=1}^n h^{j,k}(y) X_k(y), \quad y \in U,$$

so that (v) holds, by definition. Furthermore, for $1 \leq j \leq n$,

$$X_j(y) = \sum_{k=1}^n h_{j,k}(y) V_k(y),$$

so that (iii) holds for $1 \leq j \leq n$. For $n+1 \leq j \leq q$, (iii) follows from this and Lemma A.6 (vi). Since $\forall y \in B_X(x, \delta)$, $\dim \text{span}\{X_1(y), \dots, X_q(y)\} = n$, we see from (iii) that $\dim \text{span}\{V_1(y), \dots, V_n(y)\} = n$ and so (iv) follows.

It remains to prove (vi). Let $V_k^{(\alpha)} := (\phi_\alpha)_* V_k$, so that $V_k^{(\alpha)}$ is a C^1 vector field on $\phi_\alpha(U)$. By the construction of $V_k^{(\alpha)}$, $\forall u \in U$,

$$V_k^{(\alpha)}(\phi_\alpha(u)) \equiv \frac{\partial}{\partial u_k} \pmod{\left\{ \frac{\partial}{\partial u_{n+1}}, \dots, \frac{\partial}{\partial u_D} \right\}}. \quad (\text{A.5})$$

Also, by (iii) and (v), for $y \in B_X(x, \delta)$,

$$\begin{aligned} [V_j^{(\alpha)}, V_k^{(\alpha)}](\phi_\alpha(y)) &\in \text{span}\{X_1^{(\alpha)}(\phi_\alpha(y)), \dots, X_q^{(\alpha)}(\phi_\alpha(y))\} \\ &= \text{span}\{V_1^{(\alpha)}(\phi_\alpha(y)), \dots, V_n^{(\alpha)}(\phi_\alpha(y))\}. \end{aligned}$$

Combining this with (A.5), we have for $y \in B_X(x, \delta)$,

$$[V_j^{(\alpha)}, V_k^{(\alpha)}](\phi_\alpha(y)) \in \text{span}\{V_1^{(\alpha)}(\phi_\alpha(y)), \dots, V_n^{(\alpha)}(\phi_\alpha(y))\} \cap \text{span}\left\{ \frac{\partial}{\partial u_{n+1}}, \dots, \frac{\partial}{\partial u_D} \right\} = \{0\}.$$

(vi) follows, completing the proof. $\square$

LEMMA A.8. Let W and Z be C^1 vector fields on an open set $U \subseteq \mathbb{R}^d$. Then, $\forall x \in U$, $t, s \in \mathbb{R}$ such that $e^{-sZ} e^{-\tau W} e^{sZ} e^{\tau W} x$ makes sense for all $\tau \in [\min\{0, t\}, \max\{0, t\}]$, we have

$$e^{-sZ} e^{-tW} e^{sZ} e^{tW} x = x + \int_0^t \int_0^s ([W, Z](e^{-sZ} e^{-\tau W} e^{sZ} x)) \left(e^{(s-\sigma)Z} e^{\tau W}(x) \right) d\sigma d\tau,$$

where we have written $([W, Z](f))(y)$ to denote the vector field $[W, Z]$ applied to the function f , then evaluated at the point y .

Proof. This is [RS07, Lemma 4.1]. $\square$

Fix $x \in M$ and let $\alpha \in \mathcal{A}$, $U \subseteq U_\alpha$, $\delta > 0$, and $V_1, \dots, V_n$ be as in Proposition A.7. By Proposition A.7 (v), there exists $\delta_1 > 0$ such that $B_V(x, \delta_1) \subseteq B_X(x, \delta)$. For $\epsilon = \epsilon(x) > 0$ sufficiently small,¹⁸ define the map $\Phi_x : B^n(\epsilon) \rightarrow \mathfrak{M}$ by

$$\Phi_x(t_1, \dots, t_n) = e^{t_1 V_1} e^{t_2 V_2} \dots e^{t_n V_n} x.$$

Note that for $t \in B^n(\delta_1/n)$, $\Phi_x(t) \in B_V(x, \delta_1) \subseteq B_X(x, \delta) \subseteq M$.

LEMMA A.9. *For $\epsilon = \epsilon(x) > 0$ sufficiently small and for any permutation $\sigma \in S_n$,*

$$\Phi_x(t_1, \dots, t_n) = e^{t_{\sigma(1)} V_{\sigma(1)}} e^{t_{\sigma(2)} V_{\sigma(2)}} \dots e^{t_{\sigma(n)} V_{\sigma(n)}} x, \quad \forall t \in B^n(\epsilon).$$

Proof. The minor difficulty in this lemma is that $V_1, \dots, V_n$ are only known to commute on $B_X(x, \delta)$, not on a neighborhood in $\mathfrak{M}$ —since we do not yet know that $B_X(x, \delta)$ is a manifold, the lemma does not follow from standard results. We prove the lemma with $\epsilon = \delta_1/4n$. It suffices to show $\forall l \in \{1, \dots, n-1\}$,

$$e^{t_1 V_1} e^{t_2 V_2} \dots e^{t_l V_l} e^{t_{l+1} V_{l+1}} \dots e^{t_n V_n} x = e^{t_1 V_1} e^{t_2 V_2} \dots e^{t_{l-1} V_{l-1}} e^{t_{l+1} V_{l+1}} e^{t_l V_l} e^{t_{l+2} V_{l+2}} \dots e^{t_n V_n} x,$$

as the result will then follow by repeated applications of this and by symmetry in the assumptions on $V_1, \dots, V_n$. Since $e^{t_{l+2} V_{l+2}} \dots e^{t_n V_n} x \in B_V(x, \delta_1/4)$ it suffices to show $\forall(t_l, t_{l+1}) \in B^2(\epsilon)$, $y \in B_V(x, \delta_1/4)$,

$$e^{t_l V_l} e^{t_{l+1} V_{l+1}} y = e^{t_{l+1} V_{l+1}} e^{t_l V_l} y. \quad (\text{A.6})$$

Note, $\forall(t_l, t_{l+1}) \in B^2(\epsilon)$,

$$e^{-t_l V_l} e^{-t_{l+1} V_{l+1}} e^{t_l V_l} e^{t_{l+1} V_{l+1}} y \in B_V(x, \delta_1) \subseteq B_X(x, \delta).$$

Pushing this equation forward via ϕ_α gives

$$e^{-t_l V_l^{(\alpha)}} e^{-t_{l+1} V_{l+1}^{(\alpha)}} e^{t_l V_l^{(\alpha)}} e^{t_{l+1} V_{l+1}^{(\alpha)}} \phi_\alpha(y).$$

Since $[V_l^{(\alpha)}, V_{l+1}^{(\alpha)}](u) = 0$, $\forall u \in \phi_\alpha(B_V(x, \delta_1)) \subseteq \phi_\alpha(B_X(x, \delta))$, it follows from Lemma A.8 that

$$e^{-t_l V_l^{(\alpha)}} e^{-t_{l+1} V_{l+1}^{(\alpha)}} e^{t_l V_l^{(\alpha)}} e^{t_{l+1} V_{l+1}^{(\alpha)}} \phi_\alpha(y) = \phi_\alpha(y),$$

and so

$$e^{t_l V_l^{(\alpha)}} e^{t_{l+1} V_{l+1}^{(\alpha)}} \phi_\alpha(y) = e^{t_{l+1} V_{l+1}^{(\alpha)}} e^{t_l V_l^{(\alpha)}} \phi_\alpha(y).$$

(A.6) follows, completing the proof. $\square$

LEMMA A.10. *For $\epsilon = \epsilon(x) > 0$ sufficiently small,*

- (i) $\Phi_x(B^n(\epsilon)) \subseteq B_X(x, \delta)$ is an open set (and we give $\Phi_x(B^n(\epsilon))$ the subspace topology).
- (ii) $\Phi_x : B^n(\epsilon) \rightarrow \Phi_x(B^n(\epsilon))$ is a homeomorphism.
- (iii) $\Phi_x : B^n(\epsilon) \rightarrow \mathfrak{M}$ is C^2 and $d\Phi_x(u)$ has full rank (i.e., rank n), $\forall u \in B^n(\epsilon)$.
- (iv) $d\Phi_x(u) \frac{\partial}{\partial u_j} = V_j(\Phi_x(u))$.
- (v) There are C^1 vector fields $Y_1, \dots, Y_q$ on $B^n(\epsilon)$ with $\|Y_j\|_{C^1(B^n(\epsilon); \mathbb{R}^n)} < \infty$ such that $d\Phi_x(u) Y_j(u) = X_j(\Phi_x(u))$.

¹⁸ We allow $\epsilon > 0$ to shrink, as needed, throughout the argument.

Proof. We have already seen $\Phi_x(B^n(\epsilon)) \subseteq B_V(x, \delta_1) \subseteq B_X(x, \delta)$. Since $V_1, \dots, V_n$ are C^1 , standard proofs show that Φ_x is C^1 . Since $\frac{\partial}{\partial t_j} \Big|_{t=0} \Phi_x(t) = V_j(x)$ and $V_1(x), \dots, V_n(x)$ are linearly independent (Proposition A.7 (iv)) the Inverse Function Theorem shows that if $\epsilon > 0$ is sufficiently small, $\Phi_x : B^n(\epsilon) \rightarrow \mathfrak{M}$ is injective and $d\Phi_x(u)$ has full rank (i.e., rank n) $\forall u \in B^n(\epsilon)$.

By the definition of Φ_x , $\frac{\partial}{\partial t_1} \Phi_x(t) = V_1(\Phi_x(t))$, and by Lemma A.9, Φ_x is symmetric in $V_1, \dots, V_n$ and so (iv) follows for $\epsilon > 0$ sufficiently small.

Let $S \subseteq B^n(\epsilon)$ be open. We claim $\Phi_x(S) \subseteq B_X(x, \delta)$ is open. Indeed, take $\Phi_x(u) \in \Phi_x(S)$. Let $\epsilon_0 > 0$ be so small that $B^n(u, \epsilon_0) \subseteq S$. Then $\Phi_x(B^n(u, \epsilon_0)) \subseteq \Phi_x(S)$. And so $B_V(\Phi_x(u), \epsilon_0) = \Phi_x(B^n(u, \epsilon_0)) \subseteq \Phi_x(S)$.¹⁹ By Proposition A.7 (iii) $\exists \epsilon_1 > 0$ with $B_X(\Phi_x(u), \epsilon_1) \subseteq B_V(\Phi_x(u), \epsilon_0) = \Phi_x(B^n(u, \epsilon_0)) \subseteq \Phi_x(S)$. Thus, $\Phi_x(S) \subseteq B_X(x, \delta)$ is open. In particular $\Phi_x(B^n(\epsilon)) \subseteq B_X(x, \delta)$ is open. This proves (i).

Since Φ_x is an injective open map, to prove it is a homeomorphism it suffices to prove it is continuous. Let $u \in B^n(\epsilon)$ and let $S \subseteq B_X(x, \delta)$ be an open set such that $\Phi_x(u) \in S$. We wish to show that there is an open set $O \subseteq B^n(\epsilon)$, $u \in O$, $\Phi_x(O) \subseteq S$.

Take $\epsilon_0 > 0$ so small that $B_X(\Phi_x(u), \epsilon_0) \subseteq S$. Then by Proposition A.7 (iii) $\exists \epsilon_1 > 0$ with $B_V(\Phi_x(u), \epsilon_1) \subseteq B_X(\Phi_x(u), \epsilon_0) \subseteq S$. But $\Phi_x(B^n(u, \epsilon_1)) = B_V(\Phi_x(u), \epsilon_1)$; thus $O = B^n(u, \epsilon_1)$ is our desired neighborhood of u . This proves (ii).

Taking f_j^l as in Proposition A.7 (iii), and setting $Y_j(u) = \sum_{l=1}^n f_j^l \circ \Phi_x(u) \frac{\partial}{\partial u_l}$, (v) follows.

For (iii), we already know Φ_x is C^1 . That Φ_x is C^2 follows from (iv) and the fact that $V_1, \dots, V_n$ are C^1 . We have already shown $d\Phi_x(u)$ has full rank, $\forall u \in B^n(\epsilon)$. $\square$

In the previous discussion, $\epsilon > 0$ implicitly depended on x . We now make this dependence explicit and write $\epsilon_x > 0$. We consider a family of functions and open sets on M given by

$$\{(\Phi_x^{-1}, \Phi_x(B^n(\epsilon_x)))\}_{x \in M}.$$

The proof of the existence of the C^2 structure in Proposition 3.1 is completed by the next proposition.

PROPOSITION A.11. *The above maps yield a C^2 atlas on M . With this manifold structure $X_1, \dots, X_q$ are C^1 vector fields on M , and the inclusion map $M \hookrightarrow \mathfrak{M}$ is a C^2 injective immersion.*

Proof. The main point is to show that the collection of maps gives a C^2 atlas. Once this is shown, that $X_1, \dots, X_q$ are C^1 on this manifold follows from Lemma A.10 (v). That the inclusion map is a C^2 injective immersion follows from Lemma A.10 (iii).

We turn to showing the collection is a C^2 atlas. Set $W = \Phi_{x_1}(B^n(\epsilon_{x_1})) \cap \Phi_{x_2}(B^n(\epsilon_{x_2}))$. We want to show $\Phi_{x_1}^{-1} \circ \Phi_{x_2} : \Phi_{x_2}^{-1}(W) \rightarrow B^n(\epsilon_{x_1})$ is C^2 . Since $\Phi_{x_1} : B^n(\epsilon_{x_1}) \rightarrow \mathfrak{M}$ is injective, C^2 , and has injective differential (Lemma A.10 (ii) and (iii)) we have

$$\Phi_{x_1}^{-1} \circ \Phi_{x_2} \text{ is } C^2 \Leftrightarrow \Phi_{x_1} \circ \Phi_{x_1}^{-1} \circ \Phi_{x_2} \text{ is } C^2.$$

But $\Phi_{x_1} \circ \Phi_{x_1}^{-1} \circ \Phi_{x_2} = \Phi_{x_2}$ is C^2 by Lemma A.10 (iii), completing the proof. $\square$

Finally, the uniqueness of the C^2 structure in Proposition 3.1 follows immediately from the next lemma and Lemma A.1.

LEMMA A.12. *Let $\mathfrak{M}$ be a manifold and let $M \subseteq \mathfrak{M}$ be a subset. Give M any topology which is finer²⁰ than the subspace topology induced by $\mathfrak{M}$. Then, there is at most one C^2 manifold*

¹⁹ To conclude $B_V(\Phi_x(u), \epsilon_0) = \Phi_x(B^n(u, \epsilon_0))$, we have used $d\Phi_x(t) \frac{\partial}{\partial t_j} = V_j(\Phi_x(t))$ and the definition of $B_V(\Phi_x(y), \epsilon_0)$.

²⁰ Not necessarily strictly finer.

structure on M , compatible with this topology, such that the inclusion map $M \hookrightarrow \mathfrak{M}$ is an injective immersion.

Proof. Suppose there are two such C^2 structures on M ; denote the corresponding C^2 manifolds by M_1 and M_2 . We wish to show that the identity map $M_1 \rightarrow M_2$ is a C^2 diffeomorphism. Let $i_1 : M_1 \hookrightarrow \mathfrak{M}$, $i_2 : M_2 \hookrightarrow \mathfrak{M}$ be the inclusion maps (on the underlying space M , $i_1 = i_2$). Since i_1 and i_2 are assumed to be injective immersions, for all $x \in M$, there is a neighborhood $U \subseteq M$ of x such that

$$i_1|_U : M_1 \cap U \rightarrow \mathfrak{M} \cap U, \quad i_2|_U : M_2 \cap U \rightarrow \mathfrak{M} \cap U$$

are C^2 diffeomorphisms, where $\mathfrak{M} \cap U$ is given the C^2 structure as a submanifold of $\mathfrak{M}$. Hence, the identity map $U \cap M_1 \rightarrow U \cap M_2$ is a C^2 diffeomorphism. Since the identity map $M_1 \rightarrow M_2$ is a homeomorphism which is locally a C^2 diffeomorphism, we conclude that it is a global C^2 diffeomorphism, as desired. $\square$

References

- [CD06] Ph. CHARPENTIER and Y. DUPAIN. Estimates for the Bergman and Szegő projections for pseudoconvex domains of finite type with locally diagonalizable Levi form. *Publ. Mat.* (2)50 (2006), 413–446
- [CD14] Ph. CHARPENTIER and Y. DUPAIN. Extremal bases, geometrically separated domains and applications. *Algebra i Analiz* (1)26 (2014), 196–269
- [Che46] C. CHEVALLEY. Theory of Lie Groups. I. In: *Princeton Mathematical Series*, Vol. 8. Princeton University Press, Princeton (1946)
- [DP18] E. DLUGIE and A. PETERSON. On uniform large-scale volume growth for the Carnot–Carathéodory metric on unbounded model hypersurfaces in $\mathbb{C}^2$. *Involve* (1)11 (2018), 103–118
- [FS74] G. B. FOLLAND and E. M. STEIN. Estimates for the $\bar{\partial}_b$ complex and analysis on the Heisenberg group. *Comm. Pure Appl. Math.* 27 (1974), 429–522
- [FSC86] C. FEFFERMAN and A. SÁNCHEZ-CALLE. Fundamental solutions for second order subelliptic operators. *Ann. of Math.* (2) (2)124 (1986), 247–272
- [Gre15] P. GRESSMAN. Scalar oscillatory integrals in smooth spaces of homogeneous type. *Rev. Mat. Iberoam.* (1)31 (2015), 215–244
- [Gui08] V. GUILLEMIN. *Lecture notes*, <http://math.mit.edu/~arita/18.101/>, 2008, prepared by Ana Rita Pires, Accessed: 2017-05-02
- [Izz99] A. IZZO. C^r convergence of Picard’s successive approximations. *Proc. Amer. Math. Soc.* (7)127 (1999), 2059–2063
- [Lee03] J. LEE. Introduction to smooth manifolds. In: *Graduate Texts in Mathematics*, Vol. 218. Springer, New York (2003)
- [Mal69] B. MALGRANGE. Sur l’intégrabilité des structures presque-complexes. In: *Symposia Mathematica*, Vol. II (INDAM, Rome, 1968), Academic, London (1969), pp. 289–296
- [MM12] A. MONTANARI and D. MORBIDELLI. Nonsmooth Hörmander vector fields and their control balls. *Trans. Amer. Math. Soc.* (5)364 (2012), 2339–2375
- [MM13a] A. MONTANARI and D. MORBIDELLI. A Frobenius-type theorem for singular Lipschitz distributions. *J. Math. Anal. Appl.* (2)399 (2013), 692–700

- [MM13b] A. MONTANARI and D. MORBIDELLI. Step- s involutive families of vector fields, their orbits and the Poincaré inequality. *J. Math. Pures Appl. (9)* (4)99 (2013), 375–394
- [Nel59] E. NELSON. Analytic vectors. *Ann. of Math. (2)* 70 (1959), 572–615
- [NN57] A. NEWLANDER and L. NIRENBERG. Complex analytic coordinates in almost complex manifolds. *Ann. of Math. (2)* 65 (1957), 391–404
- [NRSW89] A. NAGEL, J.-P. ROSAY, E. M. STEIN, and S. WAINGER. Estimates for the Bergman and Szegő kernels in $\mathbb{C}^2$. *Ann. of Math. (2)* (1)129 (1989), 113–149
- [NS01] A. NAGEL and E. M. STEIN. Differentiable control metrics and scaled bump functions. *J. Differential Geom.* (3)57 (2001), 465–492
- [NSW85] A. NAGEL, E. M. STEIN, and S. WAINGER. Balls and metrics defined by vector fields. I. Basic properties. *Acta Math.* (1–2)155 (1985), 103–147
- [Pet14] A. PETERSON. Carnot–Carathéodory metrics in unbounded subdomains of $\mathbb{C}^2$. *Arch. Math. (Basel)* (5)102 (2014), 437–447
- [RS07] F. RAMPAZZO and H. SUSSMANN. Commutators of flow maps of nonsmooth vector fields. *J. Differential Equations* (1)232 (2007), 134–175
- [SS11] E. M. STEIN and B. STREET. Multi-parameter singular Radon transforms. *Math. Res. Lett.* (2)18 (2011), 257–277
- [SS12] E. M. STEIN and B. STREET. Multi-parameter singular Radon transforms III: Real analytic surfaces. *Adv. Math.* (4)229 (2012), 2210–2238
- [SS13] E. M. STEIN and B. STREET. Multi-parameter singular Radon transforms II: The L^p theory. *Adv. Math.* 248 (2013), 736–783
- [Ste74] P. STEFAN. Accessible sets, orbits, and foliations with singularities. *Proc. London Math. Soc. (3)* 29 (1974), 699–713
- [Ste93] E. M. STEIN. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series, Vol. 43, Princeton University Press, Princeton, NJ, 1993, With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.
- [Sto14] B. STOVALL. Uniform L^p -improving for weighted averages on curves. *Anal. PDE* (5)7 (2014), 1109–1136
- [Str11] B. STREET. Multi-parameter Carnot–Carathéodory balls and the theorem of Frobenius. *Rev. Mat. Iberoam.* (2)27 (2011), 645–732
- [Str12] B. STREET. Multi-parameter singular radon transforms I: the L^2 theory. *J. Anal. Math.* 116 (2012), 83–162
- [Str14] B. STREET. Multi-parameter singular integrals. In: *Annals of Mathematics Studies*, Vol. 189. Princeton University Press, Princeton (2014)
- [Str17] B. STREET. Sobolev spaces associated to singular and fractional Radon transforms. *Rev. Mat. Iberoam.* (2)33 (2017), 633–748
- [Str18a] B. STREET. *Coordinates Adapted to Vector Fields II: Sharp Results*, preprint [arXiv:1808.04159](https://arxiv.org/abs/1808.04159)
- [Str18b] B. STREET. *Coordinates Adapted to Vector Fields III: Real Analyticity*, preprint [arXiv:1808.04635](https://arxiv.org/abs/1808.04635)
- [Sus73] H. SUSSMANN. Orbits of families of vector fields and integrability of distributions. *Trans. Amer. Math. Soc.* 180 (1973), 171–188
- [Tri06] Hans TRIEBEL. Theory of function spaces. III. In: *Monographs in Mathematics*, Vol. 100. Birkhäuser, Basel (2006)
- [TW03] T. TAO and J. WRIGHT. L^p improving bounds for averages along curves. *J. Amer. Math. Soc.* (3)16 (2003), 605–638

BETSY STOVALL AND BRIAN STREET, Department of Mathematics, University of Wisconsin-Madison, 480 Lincoln Dr., Madison, WI, 53706, USA

`stovall@math.wisc.edu`

`street@math.wisc.edu`

Received: December 1, 2017

Revised: July 25, 2018

Accepted: August 9, 2018