



PROJECT MUSE®

---

Coordinates adapted to vector fields II: Sharp results

Brian Street

AMERICAN JOURNAL  
OF MATHEMATICS

American Journal of Mathematics, Volume 143, Number 6, December 2021,  
pp. 1791-1840 (Article)

Published by Johns Hopkins University Press

DOI: <https://doi.org/10.1353/ajm.2021.0046>



➡ For additional information about this article

<https://muse.jhu.edu/article/840132/summary>

## COORDINATES ADAPTED TO VECTOR FIELDS II: SHARP RESULTS

By 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 are  $\mathcal{C}^{s+1}$  for  $s \in (1, \infty]$ , where  $\mathcal{C}^s$  denotes the Zygmund space of order  $s$ . We give necessary and sufficient, coordinate-free conditions for the existence of such a coordinate system. Moreover, we present a quantitative study of these coordinate charts. This is the second part in a three part series of papers. The first part, joint with Stovall, addressed the same question, though the results were not sharp, and showed how such coordinate charts can be viewed as scaling maps in sub-Riemannian geometry. When viewed in this light, these results can be seen as strengthening and generalizing previous works on the quantitative theory of sub-Riemannian geometry, initiated by Nagel, Stein, and Wainger, and furthered by Tao and Wright, the author, and others. In the third part, we prove similar results concerning real analyticity.

**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  $s$ , and let  $\mathcal{C}^\infty$  denote  $C^\infty$  (for non-integer  $s$ , the Zygmund space coincides with the classical Hölder space—see Section 5.1 for more details on Zygmund spaces). In this paper, we investigate the following closely related questions for  $s \in (1, \infty]$ :

- (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 conditions for (i) and (ii), and under these conditions give a quantitative answer to (iii).

The heart of this paper is (iii); (i) and (ii) are simple consequences of our answer to (iii). The first paper in this series, joint with Stovall, [SS18] focused on a solution to (iii) which “lost one derivative”. In this paper, we take the coordinate chart developed in [SS18] as a black box, and show how to improve it to give the sharp result. The methods in [SS18] are based on ODEs, while the methods in this

---

Manuscript received December 6, 2018.

Research supported in part by NSF grants DMS-1401671 and DMS-1764265.

*American Journal of Mathematics* 143 (2021), 1791–1840. © 2021 by Johns Hopkins University Press.

paper are based on elliptic PDEs. These PDE methods were inspired by, and are closely related to, Malgrange's celebrated proof of the Newlander-Nirenberg theorem [Mal69]. In the third paper in this series, [Str21], we return to ODE methods to prove analogous results concerning real analyticity.

The coordinate charts developed in (iii) can be viewed as scaling maps in sub-Riemannian geometry. When viewed in this light, these coordinate charts can be seen as the latest results on the quantitative theory of sub-Riemannian geometry which was initiated by Nagel, Stein, and Wainger [NSW85] and C. Fefferman and Sánchez-Calle [FSC86], and furthered by many others, including Tao and Wright [TW03] and the author [Str11]. We refer the reader to [SS18] for how these charts can be viewed as scaling maps, as well as a more leisurely introduction to the questions investigated in this paper.

This paper is a continuation of the results in [SS18]. That paper gives several applications and motivations for the results described here (see, also, Remarks 2.16 and 2.17), and a more leisurely description of some of the main definitions (though we include all necessary definitions in this paper, so that the statement of the results is self-contained).

The results in this paper are a key tool in a companion paper where we study analogous questions regarding complex vector fields [Str20]. When viewed from the perspective of sub-Riemannian geometry, this companion paper allows us to create a quantitative theory of sub-Riemannian geometry which is adapted to the complex structure of a complex manifold. We call this sub-Hermitian geometry; see [Str20] for more details.

*Remark 1.1.* The results in this paper may be reminiscent of the celebrated results of DeTurck and Kazdan [DK81] regarding a coordinate system in which a Riemannian metric tensor has optimal regularity—which also used the methods introduced by Malgrange [Mal69]. However, there does not seem to be a direct relationship between our results and theirs.

*Acknowledgment.* This material is partially based upon work supported by the NSF under grant DMS-1440140 while the author was in residence at the Mathematical Sciences Research Institute in Berkeley, California, during the spring semester of 2017.

**2. Results.** In this section, we present the main results of this paper. In Section 5 (also in [SS18, Section 2]), Zygmund spaces are defined, where a distinction is made between Zygmund spaces on a subset of  $\mathbb{R}^n$ , and Zygmund spaces on a  $C^2$  manifold  $M$ . If  $\Omega \subset \mathbb{R}^n$  is a bounded, connected, open set and  $s > 0$ , we write  $\mathcal{C}^s(\Omega)$  for the classical Zygmund space of order  $s$  on  $\Omega$ ; and for a Banach space  $V$ , we write  $\mathcal{C}^s(\Omega; V)$  for the Zygmund space of order  $s$  of functions taking values in  $V$ . For a vector field  $Y = \sum_{j=1}^n a_j(t) \frac{\partial}{\partial t_j}$  on  $\Omega$ , we identify  $Y$  with the function  $(a_1, \dots, a_n) : \Omega \rightarrow \mathbb{R}^n$ , so that it makes sense to consider  $\|Y\|_{\mathcal{C}^s(\Omega; \mathbb{R}^n)}$ . We write

$\mathcal{C}^\infty(\Omega) := \bigcap_{s>0} \mathcal{C}^s(\Omega)$ , which coincides with the space of smooth functions on  $\Omega$ , all of whose derivatives are bounded on  $\Omega$ . For complete definitions and more details on  $\mathcal{C}^s(\Omega)$ , see Section 5.1.

Fix  $M$  a  $C^2$  manifold with  $C^1$  vector fields  $X_1, \dots, X_q$  on  $M$ . On  $M$ , we have the following:

- $B_X(x, \delta)$ : the sub-Riemannian ball of radius  $\delta > 0$  centered at  $x \in M$ , induced by  $X_1, \dots, X_q$ . This is defined by

$$(2.1) \quad \begin{aligned} B_X(x, \delta) \\ &:= \left\{ y \in M \mid \exists \gamma: [0, 1] \rightarrow M, \gamma(0) = x, \gamma(1) = y, \gamma \text{ is absolutely continuous,} \right. \\ &\quad \left. \gamma'(t) = \sum_{j=1}^q a_j(t) \delta X_j(\gamma(t)), a_j \in L^\infty([0, 1]), \left\| \sum_{j=1}^q |a_j|^2 \right\|_{[L^\infty]} < 1 \right\}. \end{aligned}$$

- $\rho(x, y)$ : the sub-Riemannian distance on  $M$  induced by  $X_1, \dots, X_q$ —this is the distance associated to the balls  $B_X(x, \delta)$ .

$$(2.2) \quad \rho(x, y) := \inf \{ \delta > 0 : y \in B_X(x, \delta) \}.$$

In general,  $\rho$  is merely an extended metric ( $\rho$  may take the value  $\infty$ ). However, if  $X_1, \dots, X_q$  span the tangent space at every point and  $M$  is connected, then  $\rho$  is a metric—this is the setting we are most interested in.

- $C_X^{m,s}(M)$ : the scale of Hölder spaces on  $M$ , for  $m \in \mathbb{N}$ ,  $s \in [0, 1]$ , with respect to  $X_1, \dots, X_q$ . Here, and in the rest of the paper, we use the convention  $0 \in \mathbb{N}$ .

- $\mathcal{C}_X^s(M)$ : the Zygmund space of order  $s \in (0, \infty]$  on  $M$ , with respect to  $X_1, \dots, X_q$ .

Definitions of  $C_X^{m,s}(M)$  and  $\mathcal{C}_X^s(M)$  are given in Section 5.2, and we refer the reader to [SS18] for more leisurely discussion of these spaces. We remark that the Banach spaces  $C_X^{m,s}(M)$  and  $\mathcal{C}_X^s(M)$  are defined in such a way that their norms are invariant under  $C^2$  diffeomorphisms. More precisely, if  $\Psi: N \rightarrow M$  is a  $C^2$  diffeomorphism, then

$$(2.3) \quad \|f\|_{C_X^{m,s}(M)} = \|\Psi^* f\|_{C_{\Psi^* X}^{m,s}(N)}, \quad \|f\|_{\mathcal{C}_X^s(M)} = \|\Psi^* f\|_{\mathcal{C}_{\Psi^* X}^s(N)}.$$

*Remark 2.1.* (2.3) can be interpreted as saying the norms  $\|f\|_{C_X^{m,s}(M)}$  and  $\|f\|_{\mathcal{C}_X^s(M)}$  are “coordinate-free.” In practice, this means that these norms can be computed in any  $C^2$  coordinate system, and the answer is independent of the chosen coordinate system. Moreover, it makes sense to talk about, for example,  $C_X^\infty(M) = \bigcap_m C_X^{m,0}(M)$ , even if  $M$  is merely a  $C^2$  manifold, and  $X_1, \dots, X_q$  are  $C^1$  vector fields on  $M$ .

Throughout the paper, if we say  $\|f\|_{\mathcal{C}_X^s(M)} < \infty$  we mean  $f \in \mathcal{C}_X^s(M)$  and the norm is finite, and similarly for any other function spaces.

**2.1. Qualitative results.** Let  $X_1, \dots, X_q$  be  $C^1$  vector fields on a  $C^2$  manifold  $\mathfrak{M}$ . For  $x, y \in \mathfrak{M}$ , let  $\rho(x, y)$  denote the sub-Riemannian distance associated to  $X_1, \dots, X_q$  on  $\mathfrak{M}$  defined in (2.2). Fix  $x_0 \in \mathfrak{M}$  and let  $Z := \{y \in \mathfrak{M} : \rho(x_0, y) < \infty\}$ .  $\rho$  is a metric on  $Z$ , and we give  $Z$  the topology induced by  $\rho$  (this is finer than the topology as a subspace of  $\mathfrak{M}$ , and may be strictly finer—see [SS18, Lemma A.1] for details). 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.

**PROPOSITION 2.2.** *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 2.2, see [SS18, Appendix A]. Henceforth, we assume the conditions of Proposition 2.2 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 2.3.** 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 2.4.** (The local theorem) *For  $s \in (1, \infty]$ , 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 2.5.** (ii) and (iii) of Theorem 2.4 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). Moreover, one can directly check to see if (ii) holds by computing these functions. In light of Remark 2.1, this computation can be done in any  $C^2$  coordinate system. If  $q > n$ ,  $X_1, \dots, X_q$  are linearly dependent, and the  $c_{i,j}^k$  in (iii) are not uniquely determined; (iii) only asks that there exist a choice of  $c_{i,j}^k$  satisfying the conditions in (iii). Despite this lack of uniqueness, in many applications it is more convenient to use the setting in (iii) (see, for example, the application of the quantitative results in [SS18, Section 7.1.1]).

*Remark 2.6.* Theorem 2.4 is stated for  $s \in (1, \infty]$ . It would be nice to obtain the same result for  $s \in (0, \infty]$ , however to do this with the methods of this paper, if it is even possible, would require a more technical analysis of the PDEs which arise. See Remark 6.10 for more details. Similar remarks hold for the other main results of this paper.

**THEOREM 2.7.** (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 from Theorem 2.4 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 from (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  $M$ . See Section 5.4 for formal definitions regarding  $\mathcal{C}^{s+2}$  manifolds.

*Remark 2.8.* As a corollary, we obtain results similar to Theorems 2.4 and 2.7 with the Zygmund spaces  $\mathcal{C}^{m+s}$  replaced by the easier to understand Hölder spaces  $C^{m,s}$ , with the restriction that  $s \in (0, 1)$ . For details, see Section 7.

*Remark 2.9.* The reader only wishing to understand proof of the above qualitative results, and not the more technical quantitative results, may wish to skip to the proof outline presented in Section 3.

**2.2. Quantitative results.** Theorem 2.4 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 and the quantitative control we obtain will be invariant under arbitrary  $C^2$  diffeomorphisms; see Remark 2.15. By using this quantitative control, these charts can be seen as generalized scaling maps in sub-Riemannian geometry—see [SS18, Section 7] and

Remarks 2.16 and 2.17 for more details on this and other applications. We now turn to these quantitative results, which are the heart of this paper. Because the goal is to keep track of what each constant depends on, this is somewhat technical. To ease notation, we introduce various notions of “admissible constants”; these are constants which depend only on certain parameters. While these definitions are somewhat unwieldy, they greatly simplify the statement of results and proofs throughout the paper.

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

*Definition 2.10.* For  $x \in \mathfrak{M}$ ,  $\eta > 0$ , 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

$$\mathcal{I}(n, q) := \{(i_1, i_2, \dots, i_n) : i_j \in \{1, \dots, q\}\} = \{1, \dots, q\}^n.$$

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  $x_0 \in \mathfrak{M}$  and 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  $\rho$  from (2.2). Proposition 2.2 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

$$(2.4) \quad \max_{J \in \mathcal{I}(n, q)} \left| \frac{\bigwedge X_J(x_0)}{\bigwedge X_{J_0}(x_0)} \right| \leq \zeta^{-1},$$

where  $\frac{\bigwedge X_J(x_0)}{\bigwedge X_{J_0}(x_0)}$  is defined as follows. Let  $\lambda : \bigwedge^n T_{x_0} B_X(x_0, \xi) \rightarrow \mathbb{R}$  be any nonzero linear functional; then

$$(2.5) \quad \frac{\bigwedge X_J(x_0)}{\bigwedge X_{J_0}(x_0)} := \frac{\lambda(\bigwedge X_J(x_0))}{\lambda(\bigwedge X_{J_0}(x_0))}.$$

Because  $\bigwedge^n T_{x_0} B_X(x_0, \xi)$  is one dimensional, (2.5) is independent of the choice of  $\lambda$ ; see [SS18, Section 5] for more details. Note that a  $J_0 \in \mathcal{I}(n, q)$  satisfying (2.4) always exists—one can pick  $J_0$  so that (2.4) holds with  $\zeta = 1$ ; however, it is important for some applications to have the flexibility to choose  $\zeta < 1$  (this is needed, for example, in [Str20]). 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 2.11.* Because  $X_1, \dots, X_n$  are  $C^1$ , such an  $\eta > 0$  and  $\delta_0 > 0$  always exist; see Lemma 6.12 and Remark 6.13. However, in general one can only guarantee that  $\eta, \delta_0$  are small in terms of the  $C^1$  norms of  $X_1, \dots, X_n$  in some coordinate system—and this is not a diffeomorphic invariant quantity. Thus, we state our results in terms of  $\delta_0$  and  $\eta$  to preserve the diffeomorphic invariance. See [SS18, Section 4.1] for more details on  $\eta$  and  $\delta_0$ .

*Definition 2.12.* We say  $C$  is a 0-admissible constant if  $C$  can be chosen to depend only on upper bounds for  $q, \zeta^{-1}, \xi^{-1}$ , and  $\|c_{j,k}^l\|_{C(B_{X_{J_0}}(x_0, \xi))}$ ,  $1 \leq j, k, l \leq q$ .

For the remainder of this section, fix  $s_0 > 1$ . The results which follow depend on this choice of  $s_0$ , and are stronger as  $s_0$  approaches 1.

*Definition 2.13.* 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  $\zeta, \xi, \eta$ , and  $\delta_0$ , and upper bounds for  $q$  and  $\|c_{j,k}^l\|_{\mathcal{C}_{X_{J_0}}^s(B_{X_{J_0}}(x_0, \xi))}$ ,  $1 \leq j, k, l \leq q$ . For  $s < s_0$ , we define  $\{s\}$ -admissible constants to be  $\{s_0\}$ -admissible constants.

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  $B \lesssim_{\{s\}} A$ . Similarly we define  $\lesssim_0$  and  $\approx_0$  for the same comparisons with 0-admissible constants in place of  $\{s\}$ -admissible constants.

**THEOREM 2.14.** (The quantitative theorem) *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)  $\forall y \in B_{X_{J_0}}(x_0, \chi)$ ,

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

(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  $c_{j,k}^l \in \mathcal{C}_{X_{J_0}}^{s_0}(B_{X_{J_0}}(x_0, \xi))$ , for  $1 \leq j, k, l \leq q$ . There exists a  $C^2$  map  $\Phi : B^n(1) \rightarrow B_{X_{J_0}}(x_0, \chi)$  and  $\{s_0\}$ -admissible constants  $\xi_1, \xi_2 > 0$  such that:

(d)  $\Phi(B^n(1))$  is an open subset of  $B_{X_{J_0}}(x_0, \chi)$ , and is therefore a submanifold of  $B_X(x_0, \xi)$ .

(e)  $\Phi : B^n(1) \rightarrow \Phi(B^n(1))$  is a  $C^2$  diffeomorphism.

(f)  $B_X(x_0, \xi_2) \subseteq B_{X_{J_0}}(x_0, \xi_1) \subseteq \Phi(B^n(1)) \subseteq B_X(x_0, \xi)$ .

(g)  $\Phi(0) = x_0$ .

Let  $Y_j = \Phi^* X_j$  and let  $\mathbb{M}^{n \times n}$  denote the Banach space of  $n \times n$  real matrices endowed with the operator norm. There exists an  $\{s_0\}$ -admissible  $K \geq 1$  and a matrix  $A \in \mathcal{C}^{s_0}(B^n(1); \mathbb{M}^{n \times n})$  such that:

(h)  $Y_{J_0} = K(I + A)\nabla$ , where  $\nabla$  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  $[Y_1, Y_2, \dots, Y_n]^\top$ .

(i)  $A(0) = 0$  and  $\sup_{t \in B^n(1)} \|A(t)\|_{\mathbb{M}^{n \times n}} \leq \frac{1}{2}$ .

(j) For  $s > 0$ ,  $1 \leq j \leq q$ ,

$$(2.6) \quad \|Y_j\|_{\mathcal{C}^{s+1}(B^n(1); \mathbb{R}^n)} \lesssim_{\{s\}} 1.$$

(k) We have the following equivalence of norms, for  $f \in C(B^n(1))$ ,  $s > 0$ ,

$$\|f\|_{\mathcal{C}^s(B^n(1))} \approx_{\{s-2\}} \|f\|_{\mathcal{C}_{Y_{J_0}}^s(B^n(1))} \approx_{\{s-2\}} \|f\|_{\mathcal{C}_Y^s(B^n(1))}.$$

(l) For  $f \in C(B_{X_{J_0}}(x_0, \chi))$ ,  $s > 0$ ,

$$\|f \circ \Phi\|_{\mathcal{C}^s(B^n(1))} \lesssim_{\{s-2\}} \|f\|_{\mathcal{C}_{X_{J_0}}^s(B_{X_{J_0}}(x_0, \chi))}.$$

*Remark 2.15.* The main results of this paper (including Theorem 2.14) are invariant under arbitrary  $C^2$  diffeomorphisms. This is true *quantitatively*—all of the estimates are unchanged when pushed forward under an arbitrary  $C^2$  diffeomorphism; this is a consequence of (2.3). More precisely, take  $\mathfrak{M}$  and  $X_1, \dots, X_q$  as above. Let  $N$  be another  $C^2$  manifold and let  $\Psi : M \rightarrow N$  be a  $C^2$  diffeomorphism. Then,  $X_1, \dots, X_q$  satisfy the hypotheses of Theorem 2.14 at the base point  $x_0$  if and only if  $\Psi_* X_1, \dots, \Psi_* X_q$  satisfy them at  $\Psi(x_0)$ . Moreover, admissible constants (of any kind) when defined in terms of  $X_1, \dots, X_q$  are the same as admissible constants when defined in terms of  $\Psi_* X_1, \dots, \Psi_* X_q$ . Also, if  $\Phi$  is the map guaranteed by Theorem 2.14 when applied to  $X_1, \dots, X_q$  at the point  $x_0$ , then  $\Psi \circ \Phi$  is the map

guaranteed by Theorem 2.14 when applied to  $\Psi_*X_1, \dots, \Psi_*X_q$  at the point  $\Psi(x_0)$  (as can be checked by tracing through the proof). Thus, the conclusions of Theorem 2.14 (and the other main results of this paper) remain completely unchanged when the setting is pushed forward under a  $C^2$  diffeomorphism. See [SS18] for more details.

*Remark 2.16.* As mentioned before, [SS18, Section 7] contains several applications for results like Theorems 2.14 and 2.21. Many of the applications in [SS18, Section 7] provide results in an infinitely smooth setting. By using the results in this paper (e.g., Theorem 2.14) in place of the corresponding results in [SS18] one can immediately obtain analogous results regarding a finite level of smoothness using the same proofs, which are in many ways sharp. This sharpness may be useful when studying certain non-linear PDEs defined by vector fields—where the vector fields may be defined in terms of the solution to the PDE and one does not have a priori access to smoothness estimates.

*Remark 2.17.* In Theorems 2.14 and 2.21 we have been explicit about what each constant depends on (by using the various kinds of admissible constants). In applications, what turns out to be important is what the constants do not depend on. Two simple examples of how this can work are as follows:

- We describe the setting of the foundational work of Nagel, Stein, and Wainger [NSW85]. Let  $Z_1, \dots, Z_q$  be smooth vector fields on a smooth manifold  $M$ , where each vector field  $Z_j$  is paired with a formal degree  $d_j \in [1, \infty)$ . Suppose, for  $1 \leq j, k \leq q$ ,

$$[Z_j, Z_k] = \sum_{d_l \leq d_j + d_k} c_{j,k}^l Z_l, \quad c_{j,k}^l \in \mathcal{C}_{\text{loc}}^\infty(M).$$

Set  $X_j^\delta := \delta^{d_j} Z_j$ . Then it is easy to see that  $X_1^\delta, \dots, X_q^\delta$  satisfy the hypotheses of Theorem 2.14 uniformly in  $\delta \in (0, 1]$  and uniformly as the base point  $x_0$  ranges over compact subsets of  $M$ . Thus, the conclusions of Theorem 2.14 hold uniformly in the same way; i.e., the various kinds of admissible constants can be chosen independent of  $\delta \in (0, 1]$  and  $x_0$  (as  $x_0$  ranges over a compact set). See [SS18, Section 7.1] for more details on this application. One can proceed more generally by letting the  $X_j^\delta$  depend on  $\delta$  in a more complicated way; see [SS18, Section 7.3].

- Let  $X_1, \dots, X_q$  be  $C^\infty$  vector fields on a smooth manifold  $M$ . Suppose  $[X_j, X_k] = \sum_{l=1}^q c_{j,k}^l X_l$ , where  $c_{j,k}^l \in \mathcal{C}_{\text{loc}}^\infty(M)$ . The classical Frobenius theorem applies to foliate  $M$  into leaves. This may be a singular foliation: the dimension of the leaves might not be constant. The classical proofs of the Frobenius theorem give coordinate charts which define these leaves; however these coordinate charts “blow-up” as one approaches a singular point (i.e., a point where the dimension of the leaves is not constant on any neighborhood of the point). The quantitative nature of Theorem 2.14 shows that it gives coordinate charts that avoid this blow-up in a certain sense. See [Str20] for a detailed discussion of this.

The above two examples work with  $C^\infty$  vector fields, however (as in Remark 2.16) it is straightforward to work with  $C^1$  vector fields and instead assume the hypotheses of Theorem 2.14 hold uniformly in the relevant parameters. This allows one to obtain results which are in many ways sharp in terms of regularity. We leave further details to the reader.

**2.2.1. Densities.** Let  $\chi \in (0, \xi]$  be as in Theorem 2.14. In many applications (e.g., [SS18, Section 7.1]), 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 [Gui08] (see also [Nic07] where densities are called 1-densities).

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

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

where  $\mathcal{L}_{X_j}$  denotes the Lie derivative with respect to  $X_j$ . Our goal is to understand  $\Phi^* \nu$  and  $\nu(B_X(x_0, \xi_2))$ , where  $\Phi$  and  $\xi_2$  are as in Theorem 2.14.

*Remark 2.18.* Recall, in Theorem 2.14 we fixed some  $s_0 > 1$  and all of the estimates in Theorem 2.14 were in terms of this fixed  $s_0$ . Similarly, all of the results in this section depend on this fixed choice of  $s_0$ .

*Definition 2.19.* If we say  $C$  is a  $[s_0; \nu]$ -admissible constant, it means that  $C$  is a  $\{s_0\}$ -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 2.20.* For  $s \in (0, \infty)$ , if we say  $C$  is an  $\{s; \nu\}$ -admissible constant, it means that we assume  $f_j \in \mathcal{C}_{X_{J_0}}^s(B_{X_{J_0}}(x_0, \chi))$ , and  $C$  is a  $\{s\}$ -admissible constant which is also allowed to depend on upper bounds for  $\|f_j\|_{\mathcal{C}_{X_{J_0}}^s(B_{X_{J_0}}(x_0, \chi))}$ ,  $1 \leq j \leq n$ . For  $s \leq 0$ , we define  $\{s; \nu\}$ -admissible constants to be  $[s_0; \nu]$ -admissible constants.

We write  $A \lesssim_{\{s; \nu\}} B$  for  $A \leq CB$  where  $C$  is a positive  $\{s; \nu\}$ -admissible constant. We write  $A \approx_{\{s; \nu\}} B$  for  $A \lesssim_{\{s; \nu\}} B$  and  $B \lesssim_{\{s; \nu\}} A$ . We similarly define  $\lesssim_{[s_0; \nu]}$  and  $\approx_{[s_0; \nu]}$ .

**THEOREM 2.21.** Define  $h \in C^1(B^n(1))$  by  $\Phi^* \nu = h \sigma_{\text{Leb}}$ , where  $\sigma_{\text{Leb}}$  denotes the usual Lebesgue density on  $\mathbb{R}^n$ .

(a)  $h(t) \approx_{[s_0; \nu]} \nu(X_1, \dots, X_n)(x_0)$ ,  $\forall t \in B^n(1)$ . In particular,  $h(t)$  always has the same sign, and is either never zero or always zero.

(b) For  $s > 0$ ,  $\|h\|_{\mathcal{C}^s(B^n(1))} \lesssim_{\{s-1; \nu\}} |\nu(X_1, \dots, X_n)(x_0)|$ .

**COROLLARY 2.22.** Let  $\xi_2$  be as in Theorem 2.14. Then,

$$(2.7) \quad \nu(B_{X_{J_0}}(x_0, \xi_2)) \approx_{[s_0; \nu]} \nu(B_X(x_0, \xi_2)) \approx_{[s_0; \nu]} \nu(X_1, \dots, X_n)(x_0),$$

and therefore,

$$\begin{aligned} |\nu(B_{X_{J_0}}(x_0, \xi_2))| &\approx_{[s_0; \nu]} |\nu(B_X(x_0, \xi_2))| \approx_{[s_0; \nu]} |\nu(X_1, \dots, X_n)(x_0)| \\ &\approx_0 \max_{(j_1, \dots, j_n) \in \mathcal{I}(n, q)} |\nu(X_{j_1}, \dots, X_{j_n})(x_0)|. \end{aligned}$$

**3. Outline of the proof.** The proof of Theorem 2.14 is somewhat technical. This is partially due to its quantitative nature: we keep careful track of what each constant depends on at every step. As mentioned before, this is essential for the applications we have in mind (see, e.g., Remark 2.17). In this section, we present an outline of the proof where we do not keep track of such dependencies. We hope this will help give the reader an overview of the proof before we enter into the technical details. For this section, we write  $A \lesssim B$  to mean  $A \leq CB$ , where  $C$  is a constant “which only depends on the right things;” we will make such estimates precise in the rigorous proof in later sections. To keep things simple, we outline the proof of Theorem 2.4 (iii)  $\Rightarrow$  (i) which is essentially a qualitative version of Theorem 2.14.

Fix an  $n$ -dimensional  $C^2$  manifold  $M$ , and suppose we are given  $C^1$  vector fields  $X_1, \dots, X_q$  on  $M$  which span the tangent space at every point. Fix a point  $x_0 \in M$ , and reorder the vector fields so that  $X_1(x_0), \dots, X_n(x_0)$  form a basis of  $T_{x_0}M$  ((2.4) is the assumption that  $X_1, \dots, X_n$  have nearly “maximal determinant” among all such choices). Fix  $s \in (1, \infty]$ . Our main assumption is

$$[X_i, X_j] = \sum_{k=1}^q c_{i,j}^k X_k$$

near  $x_0$ , where  $c_{i,j}^k \in \mathcal{C}_X^s$ , near  $x_0$ .

*Goal.* Our goal is to find a  $C^2$  diffeomorphism  $\Phi : B^n(1) \xrightarrow{\sim} \Phi(B^n(1))$ , where  $\Phi(B^n(1)) \subseteq M$  is an open neighborhood of  $x_0$ , such that  $\Phi^* X_j \in \mathcal{C}^{s+1}(B^n(1); \mathbb{R}^n)$ , and moreover  $\|X_j\|_{\mathcal{C}^{s+1}(B^n(1); \mathbb{R}^n)} \lesssim 1$ .

A main problem we face is that our assumptions are in terms of the diffeomorphically invariant spaces  $\mathcal{C}_X^s$ , and not in terms of standard spaces, and so we cannot initially apply standard techniques. The first step gets around this issue.

*Step 1.* The results of [SS18] (see, also, Proposition 4.1) provide a  $C^2$  diffeomorphism  $\Phi_0 : B^n(\eta_0) \xrightarrow{\sim} \Phi_0(B^n(\eta_0))$ , where  $\eta_0 \gtrsim 1$ ,  $\Phi_0(B^n(\eta_0))$  is an open neighborhood of  $x_0$ ,  $\Phi_0(0) = x_0$ , and such that if  $Y_j := \Phi_0^* X_j$ , then the following holds.

(i) We have

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = (I + A)\nabla,$$

where  $\nabla$  denotes the gradient in  $\mathbb{R}^n$  (thought of as a column vector), and  $A(t)$  is an  $n \times n$  matrix satisfying  $\|A\|_{\mathcal{C}^s(B^n(\eta_0); \mathbb{M}^{n \times n})} \lesssim 1$  and  $A(0) = 0$ .

- (ii) For  $n+1 \leq k \leq q$ ,  $Y_k = \sum_{l=1}^n b_k^l Y_l$ , where  $b_k^l \in \mathcal{C}^{s+1}(B^n(\eta_0))$ .
- (iii) For  $1 \leq j, k \leq n$ ,  $[Y_j, Y_k] = \sum_{l=1}^n \tilde{c}_{j,k}^l Y_l$ , with  $\tilde{c}_{j,k}^l \in \mathcal{C}^s(B^n(\eta_0))$ .

Step 1 achieves the goal, except with a loss of one derivative: the vector fields  $Y_1, \dots, Y_q$  are only  $\mathcal{C}^s$ , not  $\mathcal{C}^{s+1}$ . However, more is true: if all we knew was that  $Y_1, \dots, Y_q$  were  $\mathcal{C}^s$ , then we would only have (ii) and (iii) with  $s$  replaced by  $s-1$ . We will leverage this extra regularity to find a new coordinate system which completes the proof. To do this, we use methods adapted from Malgrange's work [Mal69].

*Reduction 1.* It suffices to find a  $\mathcal{C}^{s+1}$  diffeomorphism  $\Phi_2 : B^n(1) \xrightarrow{\sim} \Phi_2(B^n(1)) \subseteq B^n(\eta_0)$ , where  $\Phi_2(0) = 0$ , and such that  $\|\Phi_2^* Y_j\|_{\mathcal{C}^{s+1}} \lesssim 1$ , for  $1 \leq j \leq n$ . Indeed, given such a map  $\Phi_2$ , the goal is achieved by taking  $\Phi := \Phi_0 \circ \Phi_2$ .

*Step 2.* Fix  $\gamma_2 > 0$  small, to be chosen later. Let  $\Psi_\gamma(t) := \gamma t$ . For  $1 \leq j \leq n$ , set  $\tilde{Y}_j := \gamma \Psi_\gamma^* Y_j$ . We have

$$\begin{bmatrix} \tilde{Y}_1 \\ \tilde{Y}_2 \\ \vdots \\ \tilde{Y}_n \end{bmatrix} = (I + \tilde{A}) \nabla,$$

where  $\tilde{A}(t) = A(\gamma t)$ . Since  $A(0) = 0$ , by taking  $\gamma = \gamma(\gamma_2, \eta_0) > 0$  sufficiently small, we have

- (1)  $\|\tilde{A}\|_{\mathcal{C}^s(B^n(5); \mathbb{M}^{n \times n})} \leq \gamma_2$ .
- (2) For  $1 \leq j, k \leq n$ ,  $[\tilde{Y}_j, \tilde{Y}_k] = \sum_{l=1}^n \tilde{c}_{j,k}^l \tilde{Y}_l$ , where  $\|\tilde{c}_{j,k}^l\|_{\mathcal{C}^s(B^n(5))} \lesssim 1$  (since  $\tilde{c}_{j,k}^l = \gamma \Psi_\gamma^* \tilde{c}_{j,k}^l$ ).

*Reduction 2.* It suffices to find a  $\mathcal{C}^{s+1}$  diffeomorphism  $\Phi_1 : B^n(1) \xrightarrow{\sim} \Phi_1(B^n(1)) \subseteq B^n(5)$ , with  $\Phi_1(0) = 0$ , and such that  $\|\Phi_1^* \tilde{Y}_j\|_{\mathcal{C}^{s+1}} \lesssim 1$ ,  $1 \leq j \leq n$ . Here, we may take  $\gamma_2$  as small as we like when finding  $\Phi_1$ . We then take  $\Phi_2 := \Psi_{\gamma_1} \circ \Phi_1$  to complete the proof. This is Proposition 6.8.

*Step 3.* This step is Lemma 6.9. Fix  $\gamma_1 > 0$  small to be chosen later. By taking  $\gamma_2 = \gamma_2(\gamma_1) > 0$  sufficiently small, we find a  $\mathcal{C}^{s+1}$  diffeomorphism  $\Phi_1 : B^n(2) \xrightarrow{\sim}$

$\Phi_1(B^n(2)) \subseteq B^n(5)$ , with  $\Phi_1(0) = 0$ , such that if  $\widehat{Y}_j := \Phi_1^* \widetilde{Y}_j$ , then

$$\begin{bmatrix} \widehat{Y}_1 \\ \widehat{Y}_2 \\ \vdots \\ \widehat{Y}_n \end{bmatrix} = (I + \widehat{A})\nabla,$$

where

- If  $\widehat{A}_j$  is the  $j$ th row of  $\widehat{A}$ , then  $\sum_{j=1}^n \frac{\partial}{\partial v_j} \widehat{A}_j(v) = 0$ .
- If  $\widehat{a}_j^k$  is the  $(j, k)$  component of  $\widehat{A}$ , then  $\|\widehat{a}_j^k\|_{L^\infty} \leq \gamma_1$ .

We find  $\Phi_1$  by solving a non-linear elliptic PDE satisfied by  $\Phi_1^{-1}$ . See Lemma 6.9.

All that remains to show is that the map  $\Phi_1$  given in Step 3 satisfies the conditions of Reduction 2, provided  $\gamma_1$  is taken small enough. This is covered in Proposition 6.6. The idea is the following. We have

$$(3.1) \quad [\widehat{Y}_j, \widehat{Y}_k] = \sum_{l=1}^n \widehat{c}_{j,k}^l \widehat{Y}_l,$$

where  $\widehat{c}_{j,k}^l = \Phi_1^* \widetilde{c}_{j,k}^l \in \mathcal{C}^s(B^n(2))$ . Also, we know  $A \in \mathcal{C}^s(B^n(2))$ , but we wish to show that  $A \in \mathcal{C}^{s+1}(B^n(1))$ . To do this, note that (3.1) can be re-written as

$$\frac{\partial}{\partial t_j} \widehat{A}_k - \frac{\partial}{\partial t_k} \widehat{A}_j + \widehat{A}_j \frac{\partial}{\partial t} \widehat{A}_k - \widehat{A}_k \frac{\partial}{\partial t} \widehat{A}_j = D_{j,k},$$

where  $\frac{\partial}{\partial t} = [\frac{\partial}{\partial t_1}, \dots, \frac{\partial}{\partial t_n}]^\top$  and  $D_{j,k} \in \mathcal{C}^s$ . Combining this with  $\sum_{j=1}^n \frac{\partial}{\partial v_j} \widehat{A}_j(v) = 0$ , we see that  $\widehat{A}$  satisfies the system of equations:

$$\mathcal{E}A + \Gamma(A, \nabla A) = D,$$

where  $D \in \mathcal{C}^s$ ,  $\Gamma$  is an explicit constant coefficient bilinear form, and

$$\mathcal{E}\widehat{A} = \left( \left( \frac{\partial}{\partial t_j} \widehat{A}_k - \frac{\partial}{\partial t_k} \widehat{A}_j \right)_{1 \leq j < k \leq n}, \sum_{j=1}^n \frac{\partial}{\partial t_j} \widehat{A}_j \right).$$

By Lemma A.6,  $\mathcal{E}$  is elliptic. If  $\gamma_1$  is chosen sufficiently small, standard elliptic theory shows  $A \in \mathcal{C}^{s+1}$ , completing the proof.

*Remark 3.1.* When we turn to the rigorous proof, we present the steps in the reverse order. This is because it is much easier to make explicit the quantitative nature of each step when they are presented in the reverse order.

**4. Results from the first paper.** In this section, we describe the main result of [SS18]; namely, [SS18, Theorem 4.7]. We do not state the full result and instead state an immediate consequence of it, which is what is relevant for this paper. The setting is the same as Theorem 2.14, so that we have fixed some  $s_0 > 1$  and defined 0-admissible constants and  $\{s\}$ -admissible constants as in Definitions 2.12 and 2.13. As in Theorem 2.14 we, without loss of generality, reorder the vector fields so that  $J_0 = (1, \dots, n)$ . Set  $\eta_0 := \min\{\eta, \xi\}$  and define  $\Phi_0 : B^n(\eta_0) \rightarrow B_{X_{J_0}}(x_0, \xi)$  by

$$(4.1) \quad \Phi_0(t_1, \dots, t_n) := e^{t_1 X_1 + \dots + t_n X_n} x_0.$$

**PROPOSITION 4.1.** *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)  $\forall y \in B_{X_{J_0}}(x_0, \chi),$

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

(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 proposition, we assume  $c_{j,k}^l \in \mathcal{C}_{X_{J_0}}^{s_0}(B_{X_{J_0}}(x_0, \xi))$ , for  $1 \leq j, k, l \leq q$ . There exists an  $\{s_0\}$ -admissible constant  $\eta_1 \in (0, \eta_0]$  such that:

(d)  $\Phi_0(B^n(\eta_1))$  is an open subset of  $B_{X_{J_0}}(x_0, \chi)$  and is therefore a submanifold of  $B_X(x_0, \xi)$ .

(e)  $\Phi_0 : B^n(\eta_1) \rightarrow \Phi_0(B^n(\eta_1))$  is a  $C^2$  diffeomorphism.

Let  $Y_j := \Phi_0^* X_j$ , and write  $Y_{J_0} = (I + A)\nabla$ , where  $\nabla$  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  $[Y_1, Y_2, \dots, Y_n]^\top$ .

(f)  $A(0) = 0$  and  $\sup_{t \in B^n(\eta_1)} \|A(t)\|_{\mathbb{M}^{n \times n}} \leq \frac{1}{2}$ .

(g) For  $s > 0$ ,  $1 \leq j \leq q$ ,

$$(4.2) \quad \|Y_j\|_{\mathcal{C}^s(B^n(\eta_1); \mathbb{R}^n)} \lesssim_{\{s\}} 1.$$

(h) There exist  $b_k^l \in \mathcal{C}^{s_0+1}(B^n(\eta_1))$ ,  $n+1 \leq k \leq q$ ,  $1 \leq l \leq n$ , such that  $Y_k = \sum_{l=1}^n b_k^l Y_l$  and  $\|b_k^l\|_{\mathcal{C}^s(B^n(\eta_1))} \lesssim_{\{s-1\}} 1$ ,  $\forall s > 0$ .

(i) For  $1 \leq j, k \leq n$ ,  $[Y_j, Y_k] = \sum_{l=1}^n \tilde{c}_{j,k}^l Y_l$ , where for  $s > 0$ ,

$$\|\tilde{c}_{j,k}^l\|_{\mathcal{C}^s(B^n(\eta_1))} \lesssim_{\{s\}} 1.$$

The statement of [SS18, Theorem 4.7] uses “1-admissible constants” which we have not defined here. However, it is easy to see that 1-admissible constants are  $\{s_0\}$ -admissible constants for  $s_0 > 1$ , and so Proposition 4.1 follows from [SS18, Theorem 4.7].

*Remark 4.2.* The main difference between Proposition 4.1 and Theorem 2.14 can be seen by comparing (4.2) and (2.6): (2.6) is stronger than (4.2) by one derivative. The central point of this paper is to obtain this stronger (sharp) result.

**4.1. Densities.** We describe the results on densities from [SS18, Section 6] needed in this paper. The setting is the same as in Section 2.2.1; thus we are given a  $C^1$  density  $\nu$  on  $B_{X_{J_0}}(x_0, \xi)$  satisfying  $\mathcal{L}_{X_j}\nu = f_j\nu$ .  $[s_0; \nu]$  and  $\{s; \nu\}$ -admissible constants are defined as in that section (Definitions 2.19 and 2.20). We also use another type of admissible constant. As before, we reorder the vector fields so that  $J_0 = (1, \dots, n)$ .

*Definition 4.3.* 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  $\|f_j\|_{C(B_{X_{J_0}}(x_0, \chi))}$ ,  $1 \leq j \leq n$ . We write  $A \lesssim_{0; \nu} B$  for  $A \leq CB$ , where  $C$  is a  $0; \nu$ -admissible constant, and write  $A \approx_{0; \nu} B$  for  $A \lesssim_{0; \nu} B$  and  $B \lesssim_{0; \nu} A$ . Note that  $0; \nu$ -admissible constants are  $[s_0; \nu]$ -admissible constants.

We introduce a distinguished density on  $B_{X_{J_0}}(x_0, \chi)$  given by

$$(4.3) \quad \nu_0(Z_1, \dots, Z_n) := \left| \frac{Z_1 \wedge Z_2 \wedge \dots \wedge Z_n}{X_1 \wedge X_2 \wedge \dots \wedge X_n} \right|.$$

Note that  $X_1 \wedge X_2 \wedge \dots \wedge X_n$  is never zero on  $B_{X_{J_0}}(x_0, \chi)$  (by Proposition 4.1 (a)), so that  $\nu_0$  is defined on  $B_{X_{J_0}}(x_0, \chi)$ . It is clearly a density.

**PROPOSITION 4.4.** *Given a  $C^1$  density  $\nu$  as above, there exists  $g \in C(B_{X_{J_0}}(x_0, \chi))$  such that  $\nu = g\nu_0$  and*

(a)  $g(x) \approx_{0; \nu} g(x_0) = \nu(X_1, \dots, X_n)(x_0)$ ,  $\forall x \in B_{X_{J_0}}(x_0, \chi)$ . In particular,  $g$  always has the same sign, and is either never zero or always zero.

(b) For  $s > 0$ ,  $1 \leq j \leq n$ , we have

$$\|g\|_{\mathcal{C}_{X_{J_0}}^s(B_{X_{J_0}}(x_0, \chi))} \lesssim_{\{s-1; \nu\}} |\nu(X_1, \dots, X_n)(x_0)|.$$

*Proof.* This is an immediate consequence of [SS18, Theorem 6.5].  $\square$

**5. Function spaces.** In this section, we define the function spaces which are used in this paper as well as discuss the main properties we use. These spaces were all defined in [SS18], and we refer the reader to that paper for a more detailed discussion these spaces. As in that paper, we make a distinction between function spaces on open subsets of  $\mathbb{R}^n$  and function spaces on a  $C^2$  manifold  $M$ . Open subsets of  $\mathbb{R}^n$  have a natural smooth structure, and it makes sense to talk about the usual function spaces on these open sets. On a  $C^2$  manifold  $M$ , it does not make sense to talk about, for example,  $C^\infty$  functions. However, if we are also given  $C^1$  vector fields  $X_1, \dots, X_q$  on  $M$ , it makes sense to talk about functions which are smooth with respect to these vector fields, and that is how we proceed.

**5.1. Function spaces on Euclidean space.** In this section, we describe the standard function spaces on  $\mathbb{R}^n$  which we use. Let  $\Omega \subset \mathbb{R}^n$  be a bounded, connected, open set (we will almost always be considering the case when  $\Omega$  is a ball in  $\mathbb{R}^n$ ). We have the following classical Banach spaces of functions on  $\Omega$ :

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

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

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

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

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

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

$$(5.1) \quad \|f\|_{C^{0,s}(\Omega)} := \|f\|_{C^0(\Omega)} + \sup_{\substack{x, y \in \Omega \\ x \neq y}} |x - y|^{-s} |f(x) - f(y)|,$$

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

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

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

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

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

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

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

We set

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

It is straightforward to verify that for a ball  $B$ ,  $\mathcal{C}^\infty(B) = C^\infty(B)$ . For a Banach space  $V$ , we let  $C(\Omega; V)$ ,  $C^m(\Omega; V)$ ,  $C^{m,s}(\Omega; V)$ , and  $\mathcal{C}^s(\Omega; V)$  denote the analogous spaces of functions taking values in  $V$ . By identifying a vector field  $Y = \sum_{j=1}^n a_j \frac{\partial}{\partial t_j}$  on  $\Omega$  with the function  $(a_1, \dots, a_n) : \Omega \rightarrow \mathbb{R}^n$ , it makes sense to write, for example,  $\|Y\|_{\mathcal{C}^s(\Omega; \mathbb{R}^n)}$ .

*Remark 5.1.* The term  $\|f\|_{C^{0,s/2}(\Omega)}$  in the definition of  $\|f\|_{\mathcal{C}^s(\Omega)}$  is somewhat unusual, and is usually replaced by  $\|f\|_{C(\Omega)}$ . As is well known, if  $\Omega$  is a bounded Lipschitz domain, these two choices yield equivalent norms (this follows easily from [Tri06, Theorem 1.118 (i)]). However, the constants involved in this equivalence depend on  $\Omega$ . In this paper, we will almost always be considering the case  $\Omega = B^n(\eta)$ , for some explicit choice of  $\eta$ . Thus, the difference between these two possible definitions of  $\|f\|_{\mathcal{C}^s(\Omega)}$  will not affect any of the results in this paper. The choice we have made here is slightly more convenient for some of our purposes; see [SS18, Remark 2.1] for more comments on this.

*Definition 5.2.* For  $s \in (0, \infty]$ , we say  $f \in \mathcal{C}_{\text{loc}}^s(\Omega)$  if  $\forall x \in \Omega$ , there exists an open ball  $B \subseteq \Omega$ , centered at  $x$ , with  $f|_B \in \mathcal{C}^s(B)$ .

*Remark 5.3.* If  $\Omega$  is a bounded Lipschitz domain,  $m \in \mathbb{N}$ ,  $s \in (0, 1)$ , the spaces  $C^{m,s}(\Omega)$  and  $\mathcal{C}^{m+s}(\Omega)$  are the same—see [Tri06, Theorem 1.118 (i)]; however, if  $s \in \{0, 1\}$ , these spaces differ. As a consequence, for *any* open set  $\Omega \subseteq \mathbb{R}^n$ , for  $m \in \mathbb{N}$ ,  $s \in (0, 1)$ , we have  $\mathcal{C}_{\text{loc}}^{m+s}(\Omega)$  equals the space of functions which are locally in  $C^{m,s}$ . The space  $\mathcal{C}_{\text{loc}}^\infty(\Omega)$  equals the usual space of functions which are locally smooth on  $\Omega$ .

**5.2. Function spaces on manifolds.** Let  $X_1, \dots, X_q$  be  $C^1$  vector fields on a connected  $C^2$  manifold  $M$ . Corresponding to  $X_1, \dots, X_q$ , we have a sub-Riemannian metric given by (2.2). We use ordered multi-index notation:  $X^\alpha$ . Here,  $\alpha$  denotes a list of elements  $\{1, \dots, q\}$  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

$$\begin{aligned}\|f\|_{C_X^{0,s}(M)} &:= \|f\|_{C(M)} + \sup_{\substack{x,y \in M \\ x \neq y}} \rho(x, y)^{-s} |f(x) - f(y)|, \\ C_X^{0,s}(M) &:= \{f \in C(M) : \|f\|_{C_X^{0,s}(M)} < \infty\}.\end{aligned}$$

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

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

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 same formula as in (5.1). Given  $h > 0$ ,  $s \in (0, 1)$  define

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

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

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

It is a consequence of [SS18, 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 [SS18, Lemma 8.1]. For more details on these spaces, we refer the reader to [SS18].

*Remark 5.4.* When we write  $Vf$  for a  $C^1$  vector field  $V$  and  $f : M \rightarrow \mathbb{R}$ , we define this as  $Vf(x) := \frac{d}{dt}|_{t=0} f(e^{tV}x)$ . 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  $V_1, \dots, V_L$ , 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.

*Remark 5.5.* 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 2.2).  $B_X(x_0, \xi)$  is a metric space, with the metric  $\rho$ . 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  $X_j f(x) := \frac{d}{dt}|_{t=0} f(e^{tX_j}x)$ . Using this, we can define the spaces  $C_X^{m,s}(B_X(x_0, \xi))$  and  $\mathcal{C}_X^s(B_X(x_0, \xi))$ , and their corresponding norms, with the same formulas as above.

**5.3. Some results on function spaces.** In this section, we present some results concerning the above function spaces which we need later in the paper. Many of these results are standard and easy to prove; however a main goal of this section is to precisely state what each estimate depends on, as that is essential for our main results.

LEMMA 5.6. *For  $m \in \mathbb{N}$ ,  $s \in (0, 1)$ ,  $\eta > 0$ ,*

$$(5.2) \quad \|f\|_{C^{m,s}(B^n(\eta))} \approx \|f\|_{\mathcal{C}^{m+s}(B^n(\eta))},$$

*where the implicit constants depend on  $n$ ,  $m$ ,  $s$ , and an upper bound for  $\eta^{-1}$ . Furthermore, for  $m \in \mathbb{N}$ ,  $s \in (0, 1]$ ,  $r \in (m+s, \infty)$ ,*

$$(5.3) \quad \|f\|_{C^{m,s}(B^n(\eta))} \lesssim \|f\|_{\mathcal{C}^r(B^n(\eta))},$$

*where the implicit constant depends on  $n$ ,  $m$ ,  $s$ ,  $r$ , and an upper bound for  $\eta^{-1}$ .*

*Proof.* It suffices to prove (5.2) in the case  $m = 0$ . When  $\eta = 1$ , (5.2) (with  $m = 0$ ) follows easily from [Tri06, Theorem 1.118 (i)] (by considering the cases  $M = 1, 2$  in that theorem). For general  $\eta$ , (5.2) (with  $m = 0$ ) follows from the case  $\eta = 1$  and a simple scaling argument which we leave to the reader. (5.3) follows immediately from (5.2).  $\square$

LEMMA 5.7. *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  $m \in \mathbb{N}$ ,  $s \in (m, m+1]$ ,*

$$(5.4) \quad \|fg\|_{\mathcal{C}_X^s(M)} \leq C_{m,q} \|f\|_{\mathcal{C}_X^s(M)} \|g\|_{\mathcal{C}_X^s(M)}.$$

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

$$(5.5) \quad \|f(x)^{-1}\|_{\mathcal{C}_X^s(M)} \leq C,$$

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.* This is [SS18, Proposition 8.3].  $\square$

LEMMA 5.8. *Let  $D_1, D_2 > 0$ ,  $s_1 > 0$ ,  $s_2 \geq s_1$ ,  $s_2 > 1$ ,  $f \in \mathcal{C}^{s_1}(B^n(D_1))$ ,  $g \in \mathcal{C}^{s_2}(B^m(D_2); \mathbb{R}^n)$  with  $g(B^m(D_2)) \subseteq B^n(D_1)$ . Then,  $f \circ g \in \mathcal{C}^{s_1}(B^m(D_2))$  and  $\|f \circ g\|_{\mathcal{C}^{s_1}(B^m(D_2))} \leq C \|f\|_{\mathcal{C}^{s_1}(B^n(D_1))}$  where  $C$  can be chosen to depend only on  $s_1$ ,  $s_2$ ,  $D_1$ ,  $D_2$ ,  $m$ ,  $n$ , and an upper bound for  $\|g\|_{\mathcal{C}^{s_2}(B^m(D_2))}$ .*

Furthermore, if  $s_1 \in (0, 1)$ ,  $f$  is as above, and  $g \in C^1(B^m(D_2); \mathbb{R}^n)$  with  $g(B^m(D_2)) \subseteq B^n(D_1)$ , then  $f \circ g \in \mathcal{C}^{s_1}(B^m(D_2))$  and  $\|f \circ g\|_{\mathcal{C}^{s_1}(B^m(D_2))} \leq C \|f\|_{\mathcal{C}^{s_1}(B^n(D_1))}$  where  $C$  can be chosen to depend only on  $s_1$ ,  $D_1$ ,  $D_2$ ,  $n$ , and an upper bound for  $\|g\|_{C^1(B^m(D_2))}$ .

*Proof.* We use the notation  $A \lesssim B$  for  $A \leq CB$  where  $C$  is as in the statement of the lemma. Without loss of generality, we assume  $\|f\|_{\mathcal{C}^{s_1}(B^n(D_1))} = 1$ . We prove the first claim by induction on  $k$ , where  $s_1 \in (k, k+1]$ .

We begin with the base case  $k = 0$  so that  $s_1 \in (0, 1]$ . We use  $y$  to denote elements of  $\mathbb{R}^n$  and  $x$  to denote elements of  $\mathbb{R}^m$ . Since  $s_1 \in (0, 1]$ , we may, without loss of generality, assume  $s_2 \in (1, 2)$ ; indeed, if  $s_2 \geq 2$  we may replace  $s_2$  with  $3/2$  in the proof that follows. Since  $\|g\|_{C^1(B^m(D_2); \mathbb{R}^n)} \leq \|g\|_{\mathcal{C}^{s_2}(B^m(D_2); \mathbb{R}^n)} \lesssim 1$ , it is immediate to verify that  $\|f \circ g\|_{C^{0,s_1/2}(B^m(D_2))} \lesssim 1$ . Let  $x, h \in \mathbb{R}^m$  be such that  $x, x+h, x+2h \in B^m(D_2)$ . We wish to show

$$(5.6) \quad |f \circ g(x+2h) - 2f \circ g(x+h) + f \circ g(x)| \lesssim |h|^{s_1},$$

which will complete the proof of the base case. Define  $\gamma : [0, 2h] \rightarrow B^n(D_1)$  by  $\gamma(t) = g(x + \frac{h}{|h|}t)$ . Let  $r := (s_2 - 1)/2 \in (0, s_2 - 1)$ . We use the classical fact that  $\|g\|_{C^{1,r}(B^m(D_2); \mathbb{R}^n)} \lesssim \|g\|_{\mathcal{C}^{s_2}(B^m(D_2); \mathbb{R}^n)} \lesssim 1$  (see Lemma 5.6). Thus,  $\|\gamma\|_{C^{1,r}([0, 2|h|]; \mathbb{R}^n)} \lesssim 1$ .

Set  $\tilde{\gamma}(t) = \frac{t}{2|h|}g(x+2h) + \left(1 - \frac{t}{2|h|}\right)g(x) = \frac{t}{2|h|}\gamma(2|h|) + \left(1 - \frac{t}{2|h|}\right)\gamma(0)$ , so that  $\tilde{\gamma} : [0, 2|h|] \rightarrow B^n(D_1)$  is a line segment of length  $|g(x+2h) - g(x)| \leq$

$2|h|\|g\|_{C^1} \lesssim |h|$ . Thus, we have

$$\left| f(\tilde{\gamma}(2|h|)) - 2f(\tilde{\gamma}(|h|)) + f(\tilde{\gamma}(0)) \right| \lesssim |h|^{s_1}.$$

For  $t \in [0, 2|h|]$ , we have

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

for some  $c_1, c_2 \in [0, 2|h|]$  by the mean value theorem. Thus,

$$|\tilde{\gamma}(t) - \gamma(t)| \leq t |c_1 - c_2|^r \|\gamma\|_{C^{1,r}} \lesssim |h|^{1+r}.$$

We again use the classical fact that  $\|f\|_{C^{0,s_1/(1+r)}(B^n(D_1))} \lesssim \|f\|_{\mathcal{C}^{s_1}(B^n(D_1))} \leq 1$  (see Lemma 5.6). Thus, we have

$$\begin{aligned} |f \circ g(x+2h) - 2f \circ g(x+h) + f \circ g(x)| &= |f(\gamma(2|h|)) - 2f(\gamma(|h|)) + f(\gamma(0))| \\ &\leq |f(\tilde{\gamma}(2|h|)) - 2f(\tilde{\gamma}(|h|)) + f(\tilde{\gamma}(0))| + 2|f(\tilde{\gamma}(|h|)) - f(\gamma(|h|))| \\ &\lesssim |h|^{s_1} + |\tilde{\gamma}(|h|) - \gamma(|h|)|^{s_1/(1+r)} \|f\|_{C^{0,s_1/(1+r)}(B^n(D_1))} \lesssim |h|^{s_1}, \end{aligned}$$

completing the proof of (5.6), and therefore the proof of the base case.

Now take  $s_1 > 1$  and we assume the result for  $s_1 - 1$ . We have,

$$\|f \circ g\|_{\mathcal{C}^{s_1}(B^m(D_2))} \leq \sum_{j=1}^m \left\| \frac{\partial}{\partial x_j} (f \circ g) \right\|_{\mathcal{C}^{s_1-1}(B^m(D_2))} + \|f \circ g\|_{\mathcal{C}^{s_1-1}(B^m(D_2))}.$$

$\|f \circ g\|_{\mathcal{C}^{s_1-1}(B^m(D_2))} \lesssim 1$  by the inductive hypothesis, so it suffices to estimate  $\left\| \frac{\partial}{\partial x_j} (f \circ g) \right\|_{\mathcal{C}^{s_1-1}(B^m(D_2))}$ . We have, using Lemma 5.7,

$$\begin{aligned} \left\| \frac{\partial}{\partial x_j} (f \circ g) \right\|_{\mathcal{C}^{s_1-1}(B^m(D_2))} &\leq \sum_{l=1}^n \left\| \left( \frac{\partial f}{\partial y_l} \circ g \right) \frac{\partial g_l}{\partial x_j} \right\|_{\mathcal{C}^{s_1-1}(B^m(D_2))} \\ &\lesssim \sum_{l=1}^n \left\| \frac{\partial f}{\partial y_l} \circ g \right\|_{\mathcal{C}^{s_1-1}(B^m(D_2))} \left\| \frac{\partial g_l}{\partial x_j} \right\|_{\mathcal{C}^{s_1-1}(B^m(D_2))}. \end{aligned}$$

The inductive hypothesis shows  $\left\| \frac{\partial f}{\partial y_l} \circ g \right\|_{\mathcal{C}^{s_1-1}(B^m(D_2))} \lesssim 1$ , and

$$\left\| \frac{\partial g_l}{\partial x_j} \right\|_{\mathcal{C}^{s_1-1}(B^m(D_2))} \lesssim \|g\|_{\mathcal{C}^{s_1}(B^m(D_2))} \lesssim \|g\|_{\mathcal{C}^{s_2}(B^m(D_2))} \lesssim 1,$$

since  $s_2 \geq s_1$ . Combining the above estimates shows  $\|f \circ g\|_{\mathcal{C}^{s_1}(B^m(D_2))} \lesssim 1$ , and completes the proof of the induction.

Finally, we turn to the case when  $s_1 \in (0, 1)$  and  $g \in C^1(B^m(D_2); \mathbb{R}^n)$ . In this case, the same proof as the base case above works, by taking  $r = 0$  throughout.

Here, we use the Lemma 5.6 to see  $\|f\|_{C^{0,s_1}(B^n(D_1))} \lesssim \|f\|_{\mathcal{C}^{s_1}(B^n(D_1))}$ , for  $s_1 \in (0, 1)$ .  $\square$

LEMMA 5.9. *Fix  $s > 1$ ,  $D_1, D_2 > 0$ . Suppose  $H \in \mathcal{C}^s(B^n(D_1); \mathbb{R}^n)$  is such that  $B^n(D_2) \subseteq H(B^n(D_1))$ ,  $H : B^n(D_1) \rightarrow H(B^n(D_1))$  is a homeomorphism, and  $\inf_{t \in B^n(D_1)} |\det dH(t)| \geq c_0 > 0$ . Then,  $H^{-1} \in \mathcal{C}^s(B^n(D_2); \mathbb{R}^n)$ , with  $\|H^{-1}\|_{\mathcal{C}^s(B^n(D_2); \mathbb{R}^n)} \leq C$ , where  $C$  can be chosen to depend only on  $n, s, D_1, D_2, c_0$ , and an upper bound for  $\|H\|_{\mathcal{C}^s(B^n(D_1); \mathbb{R}^n)}$ .*

*Proof.* We use  $A \lesssim B$  for  $A \leq CB$ , where  $C$  is as in the statement of the lemma. Since  $\|H\|_{C^1(B^n(D_1); \mathbb{R}^n)} \leq \|H\|_{\mathcal{C}^s(B^n(D_1); \mathbb{R}^n)} \lesssim 1$ , we have  $H^{-1} \in C^1(B^n(D_2); \mathbb{R}^n)$  and  $\|H^{-1}\|_{C^1(B^n(D_2); \mathbb{R}^n)} \lesssim 1$ . Thus, it suffices to show

$$(5.7) \quad d(H^{-1}) \in \mathcal{C}^{s-1}(B^n(D_2); \mathbb{M}^{n \times n}) \text{ with } \|d(H^{-1})\|_{\mathcal{C}^{s-1}(B^n(D_2); \mathbb{M}^{n \times n})} \lesssim 1.$$

We use the formula

$$(5.8) \quad d(H^{-1})(t) = (dH(H^{-1}(t)))^{-1}.$$

From our hypotheses, we have  $\|dH\|_{\mathcal{C}^{s-1}(B^n(D_1); \mathbb{M}^{n \times n})} \lesssim 1$ . Since

$$\inf_{t \in B^n(D_1)} |\det dH(t)| \gtrsim 1,$$

using the cofactor representation of  $v \mapsto (dH(v))^{-1}$  and applying Lemma 5.7, we have

$$(5.9) \quad \|(dH)^{-1}\|_{\mathcal{C}^{s-1}(B^n(D_1); \mathbb{M}^{n \times n})} \lesssim 1.$$

We begin by proving (5.7) in the case  $s \in (1, 2)$ . Since

$$\|(dH)^{-1}\|_{\mathcal{C}^{s-1}(B^n(D_1); \mathbb{M}^{n \times n})} \lesssim 1$$

and  $\|H^{-1}\|_{C^1(B^n(D_1); \mathbb{R}^n)} \lesssim 1$ , it follows from Lemma 5.8 (using (5.8)) that  $\|d(H^{-1})\|_{\mathcal{C}^{s-1}(B^n(D_2); \mathbb{M}^{n \times n})} \lesssim 1$ , which completes the proof of (5.7) in this case.

We now proceed by induction. Take  $m \geq 2$  and suppose we know the lemma for  $s \in (1, m)$  and we wish to prove (5.7) for  $s \in [m, m+1]$ . Fix  $s \in [m, m+1]$ . Take  $s_1 = \frac{m+1+s}{2} - 1 \in (m-1, m)$ ; note that  $s-1 < s_1$ . By our inductive hypothesis, we have  $H^{-1} \in \mathcal{C}^{s_1}(B^n(D_2); \mathbb{R}^n)$ , with  $\|H^{-1}\|_{\mathcal{C}^{s_1}(B^n(D_2); \mathbb{R}^n)} \lesssim 1$ . Combining this with  $\|(dH)^{-1}\|_{\mathcal{C}^{s-1}(B^n(D_1); \mathbb{M}^{n \times n})} \lesssim 1$  (as shown in (5.9)) and using (5.8), Lemma 5.8 shows that  $\|d(H^{-1})\|_{\mathcal{C}^{s-1}(B^n(D_2); \mathbb{M}^{n \times n})} \lesssim 1$ , which completes the proof.  $\square$

LEMMA 5.10. *Let  $m \in \mathbb{N}$  with  $m \geq 1$ ,  $s \in (0, 1]$ , and  $\eta_1 > 0$ . For  $f \in \mathcal{C}^{m+s}(B^n(\eta_1))$  and  $\gamma \in (0, 1]$ , set  $f_\gamma(t) := f(\gamma t)$ . Then, for  $0 < \gamma \leq \min\{\frac{\eta_1}{5}, 1\}$ ,*

we have for  $f \in \mathcal{C}^{m+s}(B^n(\eta_1))$  with  $f(0) = 0$ ,

$$\|f_\gamma\|_{\mathcal{C}^{m+s}(B^n(5))} \leq \gamma 91 \|f\|_{\mathcal{C}^{m+s}(B^n(\eta_1))}.$$

*Proof.* Using  $\gamma \in (0, 1]$ , it follows immediately from the definitions that

$$\begin{aligned} (5.10) \quad & \sum_{1 \leq |\alpha| \leq m} \|\partial_x^\alpha f_\gamma\|_{\mathcal{C}^s(B^n(5))} = \sum_{1 \leq |\alpha| \leq m} \gamma^{|\alpha|} \|(\partial_x^\alpha f)(\gamma \cdot)\|_{\mathcal{C}^s(B^n(5))} \\ & \leq \sum_{1 \leq |\alpha| \leq m} \gamma^{|\alpha|} \|\partial_x^\alpha f\|_{\mathcal{C}^s(B^n(\eta_1))} \leq \gamma \|f\|_{\mathcal{C}^{m+s}(B^n(\eta_1))}. \end{aligned}$$

Since  $f_\gamma(0) = f(0) = 0$ , we have (using the Fundamental Theorem of Calculus)

$$\begin{aligned} (5.11) \quad & \|f_\gamma\|_{C^1(B^n(5))} = \|f_\gamma\|_{C^0(B^n(5))} + \sum_{|\alpha|=1} \|\partial_x^\alpha f_\gamma\|_{C^0(B^n(5))} \\ & \leq 6 \sum_{|\alpha|=1} \|\partial_x^\alpha f_\gamma\|_{C^0(B^n(5))} \leq 6\gamma \|f\|_{C^1(B^n(\eta_1))}. \end{aligned}$$

Directly from the definitions (see also [SS18, Lemma 8.1]), we have (for any ball  $B$  and any function  $g$ )

$$\|g\|_{\mathcal{C}^s(B)} \leq 5\|g\|_{C^{0,s}(B)} \leq 15\|g\|_{C^{0,1}(B)} \leq 15\|g\|_{C^1(B)} \leq 15\|g\|_{\mathcal{C}^{m+s}(B)}.$$

Thus, using (5.11), we have

$$\|f_\gamma\|_{\mathcal{C}^s(B^n(5))} \leq 15\|f_\gamma\|_{C^1(B^n(5))} \leq 90\gamma \|f\|_{C^1(B^n(\eta_1))} \leq 90\gamma \|f\|_{\mathcal{C}^{m+s}(B^n(\eta_1))}.$$

Combining this with (5.10) yields the result.  $\square$

**Remark 5.11.** For the next two results, we use the convention that for  $s \in (-1, 0]$  we set  $\mathcal{C}^s = C^{0,(s+1)/2}$  and for  $m < 0$  we set  $C^{m,s} = C^0$ , with equality of norms.

**PROPOSITION 5.12.** Fix  $\eta \in (0, 1]$ , and let  $Y_1, \dots, Y_q$  be vector fields on  $B^n(\eta)$ . We suppose  $Y_j = \sum_{j=1}^n a_j^k \frac{\partial}{\partial t_k}$  and  $\frac{\partial}{\partial t_k} = \sum_{j=1}^q b_k^j Y_j$ , for  $1 \leq j \leq q$ ,  $1 \leq k \leq n$ , where  $a_j^k \in C^1(B^n(\eta))$  and  $b_k^j \in C(B^n(\eta))$ .

• Let  $m \in \mathbb{N}$ ,  $s \in [0, 1]$ . Suppose  $a_j^k, b_k^j \in C^{m-1,s}(B^n(\eta))$ ,  $\forall j, k$ . Then,  $C^{m,s}(B^n(\eta)) = C_Y^{m,s}(B^n(\eta))$ , and

$$\|f\|_{C^{m,s}(B^n(\eta))} \approx \|f\|_{C_Y^{m,s}(B^n(\eta))},$$

where the implicit constants can be chosen to depend only on upper bounds for  $q$ ,  $m$ , and  $\|a_j^k\|_{C^{m-1,s}(B^n(\eta))}$ ,  $\|b_k^j\|_{C^{m-1,s}(B^n(\eta))}$ ,  $\forall j, k$ .

- Let  $s > 0$ . Suppose  $a_j^k, b_k^j \in \mathcal{C}^{s-1}(B^n(\eta)), \forall j, k$ . Then,  $\mathcal{C}^s(B^n(\eta)) = \mathcal{C}_Y^s(B^n(\eta))$ , and

$$\|f\|_{\mathcal{C}^s(B^n(\eta))} \approx \|f\|_{\mathcal{C}_Y^s(B^n(\eta))},$$

where the implicit constants can be chosen to depend only on  $s$  and upper bounds for  $q, \eta^{-1}$ , and  $\|a_j^k\|_{\mathcal{C}^{s-1}(B^n(\eta))}, \|b_k^j\|_{\mathcal{C}^{s-1}(B^n(\eta))}, \forall j, k$ .

*Proof.* This is [SS18, Proposition 8.12].  $\square$

**COROLLARY 5.13.** Let  $0 < \eta_1 < \eta_2$ . Let  $Y_1, \dots, Y_q$  be  $C^1$  vector fields on  $B^n(\eta_2)$  which span the tangent space to  $B^n(\eta_2)$  at every point.

- (i) For  $m \in \mathbb{N}, s \in [0, 1]$ , if  $Y_1, \dots, Y_q \in C^{m-1, s}(B^n(\eta_2); \mathbb{R}^n)$ , then

$$C^{m, s}(B^n(\eta_1)) = C_Y^{m, s}(B^n(\eta_1)).$$

- (ii) For  $s > 0$ , if  $Y_1, \dots, Y_q \in \mathcal{C}^{s-1}(B^n(\eta_2); \mathbb{R}^n)$ , then

$$\mathcal{C}^s(B^n(\eta_1)) = \mathcal{C}_Y^s(B^n(\eta_1)).$$

*Proof.* We describe the proof for (i); the proof for (ii) is similar. Since  $Y_1, \dots, Y_q \in C^{m-1, s}(B^n(\eta_2); \mathbb{R}^n)$ , we have (by definition),  $Y_j = \sum_{j=1}^n a_j^k \frac{\partial}{\partial t_k}$  with  $a_j^k \in C^{m-1, s}(B^n(\eta_2))$ . Moreover, since  $Y_1, \dots, Y_q$  span the tangent space at every point of  $B^n(\eta_2)$ , we may write  $\frac{\partial}{\partial t_k} = \sum_{j=1}^q b_k^j Y_j$ , where  $b_k^j$  is locally in  $C^{m-1, s}$ . Since  $B^n(\eta_1)$  is a relatively compact subset of  $B^n(\eta_2)$ , we see  $a_j^k, b_k^j \in C^{m-1, s}(B^n(\eta_1))$ . From here, Proposition 5.12 yields (i), completing the proof.  $\square$

**5.4. Manifolds with Zygmund regularity.** In this paper we use  $\mathcal{C}^s$  manifolds; the definition is exactly what one would expect, though a little care is needed due to the subtleties of Zygmund spaces. For example, one must define the Zygmund maps in the right way to ensure that the composition of two Zygmund maps is again a Zygmund map. For completeness, we present the relevant (standard) definitions here.

**Definition 5.14.** Let  $U_1 \subseteq \mathbb{R}^{n_1}$  and  $U_2 \subseteq \mathbb{R}^{n_2}$  be open sets. For  $s \in (0, \infty]$ , we say  $f : U_1 \rightarrow U_2$  is a  $\mathcal{C}_{\text{loc}}^s$  map if  $f \in \mathcal{C}_{\text{loc}}^s(U_1; \mathbb{R}^{n_2})$ .

**LEMMA 5.15.** Let  $U_1 \subseteq \mathbb{R}^{n_1}, U_2 \subseteq \mathbb{R}^{n_2}$ , and  $U_3 \subseteq \mathbb{R}^{n_3}$  be open sets. For  $s_1 \in (0, \infty], s_2 \geq s_1, s_2 \in (1, \infty]$ , if  $f_1 : U_1 \rightarrow U_2$  is a  $\mathcal{C}_{\text{loc}}^{s_1}$  map and  $f_2 : U_2 \rightarrow U_3$  is a  $\mathcal{C}_{\text{loc}}^{s_2}$  map, then  $f_2 \circ f_1 : U_1 \rightarrow U_3$  is a  $\mathcal{C}_{\text{loc}}^{s_1}$  map.

*Proof.* For  $s_1 = \infty$ , the result is obvious. For  $s_1 \in (0, \infty)$ , because the notion of being a  $\mathcal{C}_{\text{loc}}^s$  map is local, it suffices to check  $f_1 \circ f_2$  is in  $\mathcal{C}^{s_1}$  on sufficiently small balls. This is described in Lemma 5.8.  $\square$

LEMMA 5.16. *For  $s \in (1, \infty]$  if  $f : U_1 \rightarrow U_2$  is a  $\mathcal{C}_{\text{loc}}^s$  map which is also a  $C^1$  diffeomorphism, then  $f^{-1} : U_2 \rightarrow U_1$  is a  $\mathcal{C}_{\text{loc}}^s$  map.*

*Proof.* For  $s = \infty$ , this is standard. For  $s \in (1, \infty)$  it suffices to check  $f^{-1}$  is in  $\mathcal{C}^s$  when restricted to sufficiently small balls. This is described in Lemma 5.9.  $\square$

**Definition 5.17.** Fix  $s \in (1, \infty]$  and let  $M$  be a topological space. We say  $\{(\phi_\alpha, V_\alpha) : \alpha \in \mathcal{I}\}$  (where  $\mathcal{I}$  is some index set) is a  $\mathcal{C}^s$  atlas of dimension  $n$  if  $\{V_\alpha : \alpha \in \mathcal{I}\}$  is an open cover for  $M$ ,  $\phi_\alpha : V_\alpha \rightarrow U_\alpha$  is a homeomorphism where  $U_\alpha \subseteq \mathbb{R}^n$  is open, and  $\phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(V_\beta \cap V_\alpha) \rightarrow U_\beta$  is a  $\mathcal{C}_{\text{loc}}^s$  map.

**Definition 5.18.** For  $s \in (1, \infty]$ , a  $\mathcal{C}^s$  manifold of dimension  $n$  is a Hausdorff, paracompact topological space  $M$  endowed with a  $\mathcal{C}^s$  atlas of dimension  $n$ .

**Remark 5.19.** In this paper we assume all manifolds are paracompact. This is used in the proofs of Theorem 2.7 and Corollary 7.4 where a partition of unity is used. Otherwise, paracompactness is not used in this paper.

**Remark 5.20.** Note that an open set  $\Omega \subseteq \mathbb{R}^n$  is naturally a  $\mathcal{C}^\infty$  manifold of dimension  $n$ ; where we take the atlas consisting of a single coordinate chart (namely, the identity map  $\Omega \rightarrow \Omega$ ). We henceforth give open sets this manifold structure.

**Remark 5.21.** A  $\mathcal{C}^s$  manifold is a  $C^m$  manifold for any  $m < s$ . In light of Remark 5.3,  $\mathcal{C}^\infty$  manifolds and  $C^\infty$  manifolds are the same.

**Definition 5.22.** For  $s \in (0, \infty]$ , let  $M$  and  $N$  be  $\mathcal{C}^{s+1}$  manifolds with  $\mathcal{C}^{s+1}$  atlases  $\{(\phi_\alpha, V_\alpha)\}$  and  $\{(\psi_\beta, W_\beta)\}$ , respectively. We say  $f : M \rightarrow N$  is a  $\mathcal{C}_{\text{loc}}^{s+1}$  map if  $\psi_\beta \circ f \circ \phi_\alpha^{-1}$  is a  $\mathcal{C}_{\text{loc}}^{s+1}$  map,  $\forall \alpha, \beta$ .

LEMMA 5.23. *For  $s \in (0, \infty]$ , suppose  $M_1$ ,  $M_2$ , and  $M_3$  are  $\mathcal{C}^{s+1}$  manifolds, and  $f : M_1 \rightarrow M_2$  and  $g : M_2 \rightarrow M_3$  are  $\mathcal{C}_{\text{loc}}^{s+1}$  maps. Then,  $g \circ f : M_1 \rightarrow M_3$  is a  $\mathcal{C}_{\text{loc}}^{s+1}$  map.*

*Proof.* This follows from Lemma 5.15.  $\square$

LEMMA 5.24. *Suppose  $s \in (0, \infty]$ ,  $M_1$  and  $M_2$  are  $\mathcal{C}^{s+1}$  manifolds, and  $f : M_1 \rightarrow M_2$  is a  $\mathcal{C}_{\text{loc}}^{s+1}$  map which is also a  $C^1$  diffeomorphism. Then,  $f^{-1} : M_2 \rightarrow M_1$  is a  $\mathcal{C}_{\text{loc}}^{s+1}$  map.*

*Proof.* This follows from Lemma 5.16.  $\square$

**Definition 5.25.** Suppose  $s \in (0, \infty]$ , and  $M_1$  and  $M_2$  are  $\mathcal{C}^{s+1}$  manifolds. We say  $f : M_1 \rightarrow M_2$  is a  $\mathcal{C}^{s+1}$  diffeomorphism if  $f : M_1 \rightarrow M_2$  is a bijection and  $f : M_1 \rightarrow M_2$  and  $f^{-1} : M_2 \rightarrow M_1$  are  $\mathcal{C}_{\text{loc}}^{s+1}$  maps.

**Remark 5.26.** For  $s \in (0, \infty]$ ,  $\mathcal{C}^{s+1}$  manifolds form a category, where the morphisms are given by  $\mathcal{C}_{\text{loc}}^{s+1}$  maps. The isomorphisms in this category are exactly the  $\mathcal{C}^{s+1}$  diffeomorphisms.

For  $s \in (0, \infty]$ , a  $\mathcal{C}^{s+1}$  manifold is a  $C^1$  manifold, and it therefore makes sense to talk about vector fields on such a manifold.

*Definition 5.27.* For  $s \in (0, \infty]$  let  $M$  be a  $\mathcal{C}^{s+1}$  manifold of dimension  $n$  with  $\mathcal{C}^{s+1}$  atlas  $\{(\phi_\alpha, V_\alpha)\}$ ; here  $\phi_\alpha : V_\alpha \rightarrow U_\alpha$  is a  $\mathcal{C}^{s+1}$  diffeomorphism and  $U_\alpha \subseteq \mathbb{R}^n$  is open. We say a  $C^0$  vector field  $X$  on  $M$  is a  $\mathcal{C}^s$  vector field if  $(\phi_\alpha)_* X \in \mathcal{C}_{\text{loc}}^s(U_\alpha; \mathbb{R}^n)$ ,  $\forall \alpha$ .

**6. Proofs.** We turn to the proofs of the main results in this paper; as in the statement of Theorem 2.14, we fix some  $s_0 > 1$  throughout. The most difficult part is constructing the map  $\Phi$  from Theorem 2.14. We will construct  $\Phi$  by seeing it as a composition of two maps  $\Phi = \Phi_0 \circ \Phi_2$ , where  $\Phi_0$  is the map from Proposition 4.1 and  $\Phi_2$  is described in Section 6.1.  $\Phi_2$  itself will be constructed as a composition of two maps  $\Phi_2 = \Psi_\gamma \circ \Phi_1$ , which will be described in Section 6.6.

In the some of the sections below, we introduce new notions of  $\{s\}$ -admissible constants. We will be explicit in each section which notion we are using. These notions will be defined in such a way that the compositions described above give the proper result. For example, we prove Theorem 2.14 by reducing it to Proposition 6.3, below. Theorem 2.14 and Proposition 6.3 use different notions of  $\{s\}$ -admissible constants. However, in the application of Proposition 6.3 to prove Theorem 2.14, constants which are  $\{s\}$ -admissible in the sense of Theorem 2.14, will be  $\{s\}$ -admissible in the sense of Proposition 6.3. A similar situation occurs when we reduce Proposition 6.3 to Proposition 6.8. Thus, the various notions of  $\{s\}$ -admissible constants will seamlessly glue together to yield the main results of this paper. In each setting, once we have defined  $\{s\}$ -admissible constants, we use the notation  $A \lesssim_{\{s\}} B$  to mean  $A \leq CB$  where  $C$  is a positive  $\{s\}$ -admissible constant. And we write  $A \approx_{\{s\}} B$  for  $A \lesssim_{\{s\}} B$  and  $B \lesssim_{\{s\}} A$ .

In Section 6.1 we describe the map  $\Phi_2$ . In Section 6.2 we show how Theorem 2.14 follows by setting  $\Phi = \Phi_0 \circ \Phi_2$ . In Section 6.3 we prove the results on densities, namely Theorem 2.21 and Corollary 2.22. In Section 6.4 we state and prove a result on how to recognize the regularity of vector fields by considering their commutators. In Section 6.5 we describe and construct the map  $\Phi_1$ . In Section 6.6 we construct the map  $\Phi_2$ . Finally, in Section 6.7 we prove the qualitative results; namely Theorems 2.4 and 2.7. As mentioned in the introduction, the proofs which follow take many ideas from the work of Malgrange [Mal69].

The main idea is the following. In Proposition 4.1 we only have

$$\|Y_j\|_{\mathcal{C}^s(B^n(\eta_1); \mathbb{R}^n)} \lesssim_{\{s\}} 1,$$

but we wish to have  $\|Y_j\|_{\mathcal{C}^{s+1}(B^n(\eta_1); \mathbb{R}^n)} \lesssim_{\{s\}} 1$ . However, Proposition 4.1 gives us additional information: namely, (i), where we have  $[Y_j, Y_k] = \sum_{l=1}^n \tilde{c}_{j,k}^l Y_l$ ,  $1 \leq j, k \leq n$ , with  $\|\tilde{c}_{j,k}^l\|_{\mathcal{C}^s(B^n(\eta_1))} \lesssim_{\{s\}} 1$ . Notice, if all we knew was  $\|Y_j\|_{\mathcal{C}^s(B^n(\eta_1); \mathbb{R}^n)} \lesssim_{\{s\}} 1$  then the best we could say in general is that  $\|\tilde{c}_{j,k}^l\|_{\mathcal{C}^{s-1}(B^n(\eta_1))} \lesssim_{\{s\}} 1$ ; thus

(i) gives us additional regularity information on  $Y_1, \dots, Y_n$ . This is not enough to conclude that  $\|Y_j\|_{\mathcal{C}^{s+1}(B^n(\eta_1); \mathbb{R}^n)} \lesssim_{\{s\}} 1$ ; indeed it is easy to find two non-smooth vector fields on  $\mathbb{R}^2$ ,  $Z_1, Z_2$ , which span the tangent space at every point, such that  $[Z_1, Z_2] = 0$  (take  $Z_j = \Psi^* \frac{\partial}{\partial x_j}$  where  $\Psi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is a  $C^2$  diffeomorphism). However, as we will describe in Section 6.1, this is enough to conclude that there is a different coordinate system (denoted by  $\Phi_2$ ) in which we have  $\|\Phi_2^* Y_j\|_{\mathcal{C}^{s+1}(B^n(1); \mathbb{R}^n)} \lesssim_{\{s\}} 1$ , which will complete the proof.

**6.1.  $\Phi_2$ .** Fix  $\eta_1 > 0$  and suppose we are given vector fields  $Y_1, \dots, Y_n$  on  $B^n(\eta_1)$  of the form

$$Y = \frac{\partial}{\partial t} + A \frac{\partial}{\partial t} = (I + A)\nabla, \quad A(0) = 0.$$

Here, we are writing  $Y$  for the column vector of vector fields  $Y = [Y_1, \dots, Y_n]^\top$ ,  $\frac{\partial}{\partial t}$  is the column vector  $\frac{\partial}{\partial t} = [\frac{\partial}{\partial t_1}, \dots, \frac{\partial}{\partial t_n}]^\top$  (which we also write as  $\nabla$ ), and  $A$  is an  $n \times n$  matrix depending on  $t \in B^n(\eta_1)$ . Fix  $s_0 > 1$  and suppose  $A \in \mathcal{C}^{s_0}(B^n(\eta_1); \mathbb{M}^{n \times n})$  and

$$[Y_j, Y_k] = \sum_{l=1}^n \tilde{c}_{j,k}^l Y_l,$$

where  $\tilde{c}_{j,k}^l \in \mathcal{C}^{s_0}(B^n(\eta_1))$ .

**Definition 6.1.** For  $s \geq s_0$ , if we say  $C$  is a  $\{s\}$ -admissible constant, it means  $A \in \mathcal{C}^s(B^n(\eta_1); \mathbb{M}^{n \times n})$  and  $\tilde{c}_{j,k}^l \in \mathcal{C}^s(B^n(\eta_1))$ ,  $1 \leq j, k, l \leq n$ .  $C$  can be chosen to depend only on  $s_0$ ,  $s$  and upper bounds for  $n$ ,  $\eta_1^{-1}$ ,  $\|A\|_{\mathcal{C}^s(B^n(\eta_1); \mathbb{M}^{n \times n})}$ , and  $\|\tilde{c}_{j,k}^l\|_{\mathcal{C}^s(B^n(\eta_1))}$ . For  $s < s_0$ , we define  $\{s\}$ -admissible constants to be  $\{s_0\}$ -admissible constants.

**Remark 6.2.** In the definition of  $\{s\}$ -admissible constants, the vector fields  $Y_j$  and the functions  $\tilde{c}_{j,k}^l$  are assumed to have the same regularity. Usually, one would expect the functions  $\tilde{c}_{j,k}^l$  to be one derivative worse than the vector fields  $Y_j$ . What the following proposition shows is that one can pick a different coordinate system in which the vector fields  $Y_j$  have one more derivative of regularity, thereby achieving this expectation.

**PROPOSITION 6.3.** *There exists an  $\{s_0\}$ -admissible constant  $K \geq 1$  and a map  $\Phi_2 : B^n(1) \rightarrow B^n(\eta_1)$  such that*

(a)  $\Phi_2 \in \mathcal{C}^{s_0+1}(B^n(1); \mathbb{R}^n)$ , and

$$\|\Phi_2\|_{\mathcal{C}^{s+1}(B^n(1); \mathbb{R}^n)} \lesssim_{\{s\}} 1, \quad \forall s > 0.$$

(b)  $\Phi_2(0) = 0$ ,  $d\Phi_2(0) = K^{-1}I$ .

(c)  $\Phi_2(B^n(1)) \subseteq B^n(\eta_1)$  is open and  $\Phi_2 : B^n(1) \rightarrow \Phi_2(B^n(1))$  is a  $\mathcal{C}^{s_0+1}$  diffeomorphism.

Let  $\widehat{Y}_j = \Phi_2^* Y_j$ . Then,

$$(d) \quad \widehat{Y} = K(I + \widehat{A})\nabla, \text{ and } \widehat{A}(0) = 0.$$

$$(e) \quad \sup_{u \in B^n(1)} \|\widehat{A}(u)\|_{\mathbb{M}^{n \times n}} \leq \frac{1}{2}.$$

$$(f) \quad \|\widehat{Y}_j\|_{\mathcal{C}^{s+1}(B^n(1); \mathbb{R}^n)} \lesssim_{\{s\}} 1, \text{ for } s > 0, 1 \leq j \leq n.$$

We defer the proof of Proposition 6.3 to Section 6.6.

**6.2. Proof of Theorem 2.14.** In this section, we prove Theorem 2.14 by combining Propositions 4.1 and 6.3. We take the same setting as in Theorem 2.14, and define 0-admissible and  $\{s\}$ -admissible constants as in Definitions 2.12 and 2.13. Take  $\Phi_0, Y_1, \dots, Y_q, A, \eta_1$ , and  $\chi$  be as in Proposition 4.1, so that  $\Phi_0 : B^n(\eta_1) \rightarrow B_{X_{J_0}}(x_0, \chi)$ . Note that (4.2) implies  $\|A\|_{\mathcal{C}^s(B^n(\eta_1); \mathbb{M}^{n \times n})} \lesssim_{\{s\}} 1$ . Hence, using Proposition 4.1 (f), (g), and (i), we see that Proposition 6.3 applies to  $Y_1, \dots, Y_n$  (with this choice of  $\eta_1$ ), and every constant which is  $\{s\}$ -admissible in the sense of Proposition 6.3 is  $\{s\}$ -admissible in the sense of this section. Thus we obtain a map  $\Phi_2 : B^n(1) \rightarrow B^n(\eta_1)$  as in Proposition 6.3. Let  $K, \widehat{A}$ , and  $\widehat{Y}_1, \dots, \widehat{Y}_n$  be as in that proposition. Notationally, we prove Theorem 2.14 with  $\widehat{Y}$  in place of  $Y$  and  $\widehat{A}$  in place of  $A$ .

With  $\chi \in (0, \xi]$  as in Proposition 4.1, Theorem 2.14 (a), (b), and (c) follow immediately from Proposition 4.1 (a), (b), and (c). Set  $\Phi = \Phi_0 \circ \Phi_2 : B^n(1) \rightarrow B_{X_{J_0}}(x_0, \chi)$ .

By Proposition 4.1 (d) and (e),  $\Phi_0$  takes open subsets of  $B^n(\eta_1)$  to open subsets of  $B_{X_{J_0}}(x_0, \chi)$ . By Proposition 6.3 (c),  $\Phi_2(B^n(1))$  is open in  $B^n(\eta_1)$ . Theorem 2.14 (d) follows. Theorem 2.14 (e) follows by combining Proposition 4.1 (e) and Proposition 6.3 (c).

By the definition of  $\Phi_0$ , (4.1), we have  $\Phi_0(0) = x_0$ . By Proposition 6.3 (b), we have  $\Phi_2(0) = 0$ . Hence,  $\Phi(0) = x_0$ , proving Theorem 2.14 (g). The existence of  $\xi_1$  as in Theorem 2.14 (f) follows just as in [SS18, Lemma 9.23], while the existence of  $\xi_2$  follows from [SS18, Lemma 9.35].

For  $1 \leq j \leq n$ , we have  $\Phi^* X_j = \Phi_2^* \Phi_0^* X_j = \Phi_2^* Y_j = \widehat{Y}_j$ . For  $n+1 \leq j \leq q$ , we define  $\widehat{Y}_j := \Phi^* X_j$ . Proposition 6.3 (d) and (e) shows  $\widehat{Y}_{J_0} = K(I + \widehat{A})\nabla$  and proves Theorem 2.14 (h) and (i).

Proposition 6.3 (f) proves Theorem 2.14 (j) for  $1 \leq j \leq n$ . For  $n+1 \leq j \leq q$ , we proceed as follows. Let  $b_j^l$  be as in Proposition 4.1 (h). Then, we have

$$(6.1) \quad \widehat{Y}_j = \Phi_2^* Y_j = \sum_{k=1}^n \Phi_2^*(b_j^k Y_k) = \sum_{k=1}^n (b_j^k \circ \Phi_2) \widehat{Y}_k.$$

We have already shown

$$(6.2) \quad \|\widehat{Y}_k\|_{\mathcal{C}^{s+1}(B^n(1); \mathbb{R}^n)} \lesssim_{\{s\}} 1, \quad 1 \leq k \leq n.$$

Since  $\|b_j^k\|_{\mathcal{C}^{s+1}(B^n(\eta_1))} \lesssim_{\{s\}} 1$  by Proposition 4.1 (h) and  $\|\Phi_2\|_{\mathcal{C}^{s+1}(B^n(1);\mathbb{R}^n)} \lesssim_{\{s\}} 1$  by Proposition 6.3 (a), we have  $\|b_j^k \circ \Phi_2\|_{\mathcal{C}^{s+1}(B^n(1))} \lesssim_{\{s\}} 1$  for  $s > 0$  (see Lemma 5.8). Combining this with (6.1) and (6.2) completes the proof of Theorem 2.14 (j).

Notice that Theorem 2.14 (j) (which we have already shown) implies  $\|\widehat{A}\|_{\mathcal{C}^{s+1}(B^n(1);\mathbb{M}^{n \times n})} \lesssim_{\{s\}} 1$ . We have  $\widehat{Y}_{J_0} = K(I + \widehat{A})\nabla$ . Since  $\|\widehat{A}(u)\|_{\mathbb{M}^{n \times n}} \leq \frac{1}{2}$ ,  $\forall u \in B^n(1)$ ,  $(I + \widehat{A}(u))$  is invertible for all  $u \in B^n(1)$  and we have  $\|(I + \widehat{A})^{-1}\|_{\mathcal{C}^{s+1}(B^n(1))} \lesssim_{\{s\}} 1$  (this uses Lemma 5.7 and the cofactor representation of  $(I + \widehat{A})^{-1}$ ). Hence,  $\nabla = K^{-1}(I + \widehat{A})^{-1}\widehat{Y}_{J_0}$ . I.e., for each  $1 \leq j \leq n$ ,  $\frac{\partial}{\partial t_j}$  can be written as a linear combination, with coefficients in  $\mathcal{C}^{s+1}(B^n(1))$  of  $\widehat{Y}_1, \dots, \widehat{Y}_n$ , and the  $\mathcal{C}^{s+1}$  norms of the coefficients are  $\lesssim_{\{s\}} 1$ . Combining this with Theorem 2.14 (j), Proposition 5.12 applies to prove Theorem 2.14 (k).

For Theorem 2.14 (l), we already know by Theorem 2.14 (k) that

$$\|f \circ \Phi\|_{\mathcal{C}^s(B^n(1))} \approx_{\{s-2\}} \|f \circ \Phi\|_{\mathcal{C}_{\widehat{Y}_{J_0}}^s(B^n(1))}.$$

That  $\|f \circ \Phi\|_{\mathcal{C}_{\widehat{Y}_{J_0}}^s(B^n(1))} \leq \|f\|_{\mathcal{C}_{X_{J_0}}^s(B_{X_{J_0}}(x_0, \chi))}$  follows from [SS18, Proposition 8.6]; Theorem 2.14 (l) follows.

**6.3. Densities.** In this section, we prove Theorem 2.21 and Corollary 2.22. We take the setting of Theorem 2.21 and therefore we have a  $C^1$  density  $\nu$  and a notion of  $\{s; \nu\}$ -admissible constants, as in Definition 2.20. We let  $\Phi, Y_1, \dots, Y_q, K$ , and  $A$  be as in Theorem 2.14, and we let  $\nu_0$  be as in (4.3).

**LEMMA 6.4.** *Define  $h_0$  by  $\Phi^*\nu_0 = h_0\sigma_{\text{Leb}}$ . Then,  $h_0 = \det(K(I + A))^{-1}$ . In particular,  $h_0(t) \approx_{\{s_0\}} 1$ ,  $\forall t \in B^n(1)$ , and*

$$(6.3) \quad \|h_0\|_{\mathcal{C}^s(B^n(1))} \lesssim_{\{s-1\}} 1, \quad s > 0.$$

*Proof.* Because  $\sup_{t \in B^n(1)} \|A(t)\|_{\mathbb{M}^{n \times n}} \leq \frac{1}{2}$  and  $K \approx_{\{s_0\}} 1$  by Theorem 2.14, we have  $|\det(K(I + A))^{-1}| = \det(K(I + A))^{-1}$ , and  $\det(K(I + A))^{-1} \approx_{\{s_0\}} 1$ . Using that  $\Phi_*Y_j = X_j$ ,

$$\begin{aligned} h_0(t) &= (\Phi^*\nu_0)(t) \left( \frac{\partial}{\partial t_1}, \frac{\partial}{\partial t_2}, \dots, \frac{\partial}{\partial t_n} \right) \\ &= (\Phi^*\nu_0)(t) \left( (K(I + A(t)))^{-1}Y_1(t), \dots, (K(I + A(t)))^{-1}Y_n(t) \right) \\ &= \left| \det(K(I + A(t)))^{-1} \right| (\Phi^*\nu_0)(t) (Y_1(t), \dots, Y_n(t)) \\ &= \det(K(I + A(t)))^{-1} \nu_0(\Phi(t)) (X_1(\Phi(t)), \dots, X_n(\Phi(t))) \\ &= \det(K(I + A(t)))^{-1}. \end{aligned}$$

This proves  $h_0 = \det(K(I + A))^{-1}$  and therefore  $h_0(t) \approx_{\{s_0\}} 1$ . Theorem 2.14 (j) implies  $\|A\|_{\mathcal{C}^s(B^n(1); \mathbb{M}^{n \times n})} \lesssim_{\{s-1\}} 1$ ; (6.3) follows from this using Lemma 5.7, completing the proof.  $\square$

*Proof of Theorem 2.21.* Let  $g$  be as in Proposition 4.4 so that  $\nu = g\nu_0$ . Hence,  $h\sigma_{\text{Leb}} = \Phi^*\nu = \Phi^*g\nu_0 = (g \circ \Phi)h_0\sigma_{\text{Leb}}$ , where  $h_0$  is as in Lemma 6.4. Thus,  $h = (g \circ \Phi)h_0$ . Proposition 4.4(a) implies  $g \circ \Phi(t) \approx_{[s_0; \nu]} \nu(X_1, \dots, X_n)(x_0)$  and Lemma 6.4 shows  $h_0(t) \approx_{\{s_0\}} 1$ . (a) follows.

Theorem 2.14(l) combined with Proposition 4.4(b) shows  $\|g \circ \Phi\|_{\mathcal{C}^s(B^n(1))} \lesssim_{\{s-1; \nu\}} 1$ . Combining this with (6.3) and the formula  $h = (g \circ \Phi)h_0$ , and using Lemma 5.7, proves (b) and completes the proof.  $\square$

To prove Corollary 2.22, we introduce a corollary of Theorem 2.14.

**COROLLARY 6.5.** *Let  $\Phi$ ,  $\xi_1$ , and  $\xi_2$  be as in Theorem 2.14. Then, there exist  $\{s_0\}$ -admissible constants  $0 < \xi_4 \leq \xi_3 \leq \xi_2$  and a map  $\widehat{\Phi} : B^n(1) \rightarrow B_{X_{J_0}}(x_0, \xi_2)$  which satisfies all the same estimates as  $\Phi$  so that*

$$\begin{aligned} B_X(x_0, \xi_4) &\subseteq B_{X_{J_0}}(x_0, \xi_3) \subseteq \widehat{\Phi}(B^n(1)) \subseteq B_{X_{J_0}}(x_0, \xi_2) \subseteq B_X(x_0, \xi_2) \\ &\subseteq B_{X_{J_0}}(x_0, \xi_1) \subseteq \Phi(B^n(1)) \subseteq B_{X_{J_0}}(x_0, \chi) \subseteq B_{X_{J_0}}(x_0, \xi). \end{aligned}$$

*Proof.* After obtaining  $\xi_1$ ,  $\xi_2$ , and  $\Phi$  from Theorem 2.14, we apply Theorem 2.14 again with  $\xi$  replaced by  $\xi_2$ , to yield the map  $\widehat{\Phi}$  and  $\{s_0\}$ -admissible constants  $\xi_3$  and  $\xi_4$  as above.  $\square$

*Proof of Corollary 2.22.* Using Theorem 2.21(a), we have

$$\begin{aligned} \nu(\Phi(B^n(1))) &= \int_{\Phi(B^n(1))} \nu = \int_{B^n(1)} \Phi^*\nu \\ &= \int_{B^n(1)} h(t) dt \approx_{[s_0; \nu]} \nu(X_1, \dots, X_n)(x_0), \end{aligned}$$

and we have the same estimate for  $\Phi$  replaced by  $\widehat{\Phi}$ , where  $\widehat{\Phi}$  is as in Corollary 6.5. Since

$$\widehat{\Phi}(B^n(1)) \subseteq B_{X_{J_0}}(x_0, \xi_2) \subseteq B_X(x_0, \xi_2) \subseteq \Phi(B^n(1)),$$

and since  $h(t)$  always has the same sign (by Theorem 2.21(a)), (2.7) follows.

To complete the proof, we need to show

$$(6.4) \quad |\nu(X_1, \dots, X_n)(x_0)| \approx_0 \max_{(j_1, \dots, j_n) \in \mathcal{I}(n, q)} |\nu(X_{j_1}, \dots, X_{j_n})(x_0)|.$$

However, either both sides of this equation equal 0, or Proposition 4.4 shows

$$\begin{aligned} \frac{|\nu(X_{j_1}, \dots, X_{j_n})(x_0)|}{|\nu(X_1, \dots, X_n)(x_0)|} &= \frac{|\nu_0(X_{j_1}, \dots, X_{j_n})(x_0)|}{|\nu_0(X_1, \dots, X_n)(x_0)|} \\ &= \left| \frac{X_{j_1}(x_0) \wedge \dots \wedge X_{j_n}(x_0)}{X_1(x_0) \wedge \dots \wedge X_n(x_0)} \right| \leq \zeta^{-1} \lesssim_0 1, \end{aligned}$$

where we have used the definition of  $\zeta$  (see (2.4)). Since the left-hand side of (6.4) is  $\leq$  the right-hand side, this completes the proof.  $\square$

**6.4. A regularity result.** Let  $Y_1, \dots, Y_n$  be vector fields on  $B^n(2)$ . Using the vector notation from Section 6.1, write

$$Y = \frac{\partial}{\partial t} + A \frac{\partial}{\partial t},$$

where  $A : B^n(2) \rightarrow \mathbb{M}^{n \times n}$ . Let  $a_j^k$  denote the  $(j, k)$  component of  $A$ , and define  $A_j = [a_j^1, \dots, a_j^n]$ ; i.e.,  $A_j$  is the  $j$ th row of  $A$ . We have

$$Y_j = \frac{\partial}{\partial t_j} + A_j \frac{\partial}{\partial t}.$$

Suppose

$$(6.5) \quad [Y_j, Y_k] = \sum_{l=1}^n c_{j,k}^l Y_l$$

and

$$(6.6) \quad \sum_{j=1}^n \frac{\partial}{\partial t_j} A_j = 0.$$

**PROPOSITION 6.6.** *In the above setting, there exists  $\gamma_1 = \gamma_1(n) > 0$  (depending only on  $n$ ) such that the following holds. If  $s > 1$  is such that  $c_{j,k}^l, a_j^k \in \mathcal{C}^s(B^n(2))$ ,  $\forall j, k, l$ , and  $\|a_j^k\|_{L^\infty(B^n(2))} \leq \gamma_1$ ,  $\forall j, k$ , then  $a_j^k \in \mathcal{C}^{s+1}(B^n(1))$  and*

$$\max_{j,k} \|a_j^k\|_{\mathcal{C}^{s+1}(B^n(1))} \leq D_{n,s},$$

where  $D_{n,s}$  can be chosen to depend only on  $s$ , and upper bounds for  $n$ ,  $\|a_j^k\|_{\mathcal{C}^s(B^n(2))}$ , and  $\|c_{j,k}^l\|_{\mathcal{C}^s(B^n(2))}$  (for all  $j, k, l$ ).

*Proof.* Set  $C_{j,k} = [c_{j,k}^1, \dots, c_{j,k}^n]$ . Then (6.5) can be rewritten as

$$\frac{\partial}{\partial t_j} A_k - \frac{\partial}{\partial t_k} A_j + A_j \frac{\partial}{\partial t} A_k - A_k \frac{\partial}{\partial t} A_j = C_{j,k}(I + A).$$

Combining this with (6.6) shows that  $A$  satisfies the following system of equations:

$$\mathcal{E}A + \Gamma(A, \nabla A) = \widehat{C},$$

where

$$\mathcal{E}A = \left( \left( \frac{\partial}{\partial t_j} A_k - \frac{\partial}{\partial t_k} A_j \right)_{1 \leq j < k \leq n}, \sum_{j=1}^n \frac{\partial}{\partial t_j} A_j \right),$$

$\Gamma$  is a constant coefficient bilinear form, depending only on  $n$ , and  $\widehat{C} = ((C_{j,k}(I + A))_{1 \leq j < k \leq n}, 0)$ .

By Lemma A.6,  $\mathcal{E}$  is elliptic. Also,  $\|\widehat{C}\|_{\mathcal{C}^s} \leq D_{n,s}$ , where  $D_{n,s}$  is as in the statement of the proposition (see Lemma 5.7). From here, the result follows from Proposition A.3 (taking  $s_1 = s - 1$  and  $s_2 = s$  in that proposition).  $\square$

**6.5.**  $\Phi_1$ . Fix  $s_0 > 1$ . Let  $Y_1, \dots, Y_n$  be  $\mathcal{C}^{s_0}$  vector fields on  $B^n(5)$ . Using the matrix notation of Section 6.1, we assume  $Y_1, \dots, Y_n$  have the form

$$Y = \frac{\partial}{\partial t} + A \frac{\partial}{\partial t}, \quad A(0) = 0,$$

where  $A : B^n(5) \rightarrow \mathbb{M}^{n \times n}$ . We assume

$$[Y_j, Y_k] = \sum_{l=1}^n c_{j,k}^l Y_l.$$

**Definition 6.7.** For  $s \geq s_0$ , if we say  $C$  is a  $\{s\}$ -admissible constant it means that  $A \in \mathcal{C}^s(B^n(5); \mathbb{M}^{n \times n})$  and  $c_{j,k} \in \mathcal{C}^s(B^n(5))$ ,  $1 \leq j, k, l \leq n$ .  $C$  can be chosen to depend only on  $s$ ,  $s_0$ ,  $n$ , and upper bounds for  $\|A\|_{\mathcal{C}^s(B^n(5); \mathbb{M}^{n \times n})}$  and  $\|c_{j,k}^l\|_{\mathcal{C}^s(B^n(5))}$ ,  $1 \leq j, k, l \leq n$ . For  $s < s_0$ , we define  $\{s\}$ -admissible constants to be  $\{s_0\}$ -admissible constants.

**PROPOSITION 6.8.** *There exists  $\gamma_2 = \gamma_2(n, s_0) > 0$  ( $\gamma_2$  depending only on  $n$  and  $s_0$ ) such that if  $\|A\|_{\mathcal{C}^{s_0}(B^n(5); \mathbb{M}^{n \times n})} \leq \gamma_2$  then there exists  $\Phi_1 : B^n(1) \rightarrow B^n(5)$  such that:*

- (a)  $\Phi_1 \in \mathcal{C}^{s_0+1}(B^n(1); \mathbb{R}^n)$  and  $\|\Phi_1\|_{\mathcal{C}^{s_0+1}(B^n(1); \mathbb{R}^n)} \leq D_{n,s_0}$ , where  $D_{n,s_0}$  depends only on  $n$  and  $s_0$ .
- (b)  $\|\Phi_1\|_{\mathcal{C}^{s+1}(B^n(1))} \lesssim_{\{s\}} 1, \forall s > 0$ .
- (c)  $\Phi_1(0) = 0$  and  $d\Phi_1(0) = I$ .
- (d)  $\Phi_1(B^n(1)) \subseteq B^n(5)$  is open.
- (e)  $\Phi_1 : B^n(1) \rightarrow \Phi_1(B^n(1))$  is a  $\mathcal{C}^{s_0+1}$  diffeomorphism.

Let  $\widehat{Y}_j := \Phi_1^* Y_j$ , then

$$\widehat{Y} = \frac{\partial}{\partial t} + \widehat{A} \frac{\partial}{\partial t},$$

where

- (f)  $\widehat{A}(0) = 0$  and  $\sup_{u \in B^n(1)} \|\widehat{A}(u)\|_{\mathbb{M}^{n \times n}} \leq \frac{1}{2}$ .
- (g)  $\|\widehat{A}\|_{\mathcal{C}^{s+1}(B^n(1); \mathbb{M}^{n \times n})} \lesssim_{\{s\}} 1, s > 0$ .
- (h)  $\|\widehat{Y}_j\|_{\mathcal{C}^{s+1}(B^n(1); \mathbb{R}^n)} \lesssim_{\{s\}} 1, s > 0$ .

The rest of this section is devoted to the proof of Proposition 6.8.

LEMMA 6.9. *Fix  $\sigma, \gamma_1 > 0$ . There exists  $\gamma_2 = \gamma_2(n, s_0, \sigma, \gamma_1) > 0$  ( $\gamma_2$  depending only on  $n, s_0, \sigma$ , and  $\gamma_1$ ) such that if  $\|A\|_{\mathcal{C}^{s_0}(B^n(5); \mathbb{M}^{n \times n})} \leq \gamma_2$  then there exists  $H \in \mathcal{C}^{s_0+1}(B^n(4); \mathbb{R}^n)$  of the form  $H(t) = t + R(t)$  where*

(a)  $H(B^n(4)) \subseteq \mathbb{R}^n$  is open and  $H : B^n(4) \rightarrow H(B^n(4))$  is a  $\mathcal{C}^{s_0+1}$  diffeomorphism.

- (b)  $R(0) = 0$  and  $dR(0) = 0$ .
- (c)  $R \in \mathcal{C}^{s_0+1}(B^n(4); \mathbb{R}^n)$  with  $\|R\|_{\mathcal{C}^{s_0+1}(B^n(4); \mathbb{R}^n)} \leq \sigma$ .
- (d)  $\|R\|_{\mathcal{C}^{s+1}(B^n(3); \mathbb{R}^n)} \lesssim_{\{s\}} 1$  for all  $s > 0$ .

Moreover, let  $\widehat{Y}_j = H_* Y_j$ . Then  $\widehat{Y} = \frac{\partial}{\partial v} + \widehat{A} \frac{\partial}{\partial v}$  and

- (e) If  $\widehat{A}_j$  is the  $j$ th row of  $\widehat{A}$ , then  $\sum_{j=1}^n \frac{\partial}{\partial v_j} \widehat{A}_j(v) = 0$  for  $v \in H(B^n(4))$ .
- (f) If  $\widehat{a}_j^k$  is the  $(j, k)$  component of  $\widehat{A}$ , then  $\|\widehat{a}_j^k\|_{L^\infty(H(B^n(4)))} \leq \gamma_1$ .

*Proof.* If  $\sigma > 0$  is sufficiently small, depending only on  $s_0$  and  $n$ , and if (c) holds, the Inverse Function Theorem implies (a). Thus, without loss of generality, we shrink  $\sigma > 0$  so that (a) holds. (d) for  $s < s_0$  follows from the result for  $s = s_0$  (by the definition of  $\{s\}$ -admissible constants). Thus it suffices to prove (d) for  $s \geq s_0$ .

To begin, let  $R \in \mathcal{C}^{s_0+1}(B^n(4); \mathbb{R}^n)$  be any function satisfying  $R(0) = 0$ ,  $dR(0) = 0$ , and  $\|R\|_{\mathcal{C}^{s_0+1}(B^n(4); \mathbb{R}^n)} \leq \sigma$  (we will later specialize to a specific choice of  $R$ ). To emphasize the dependence of  $H$  on  $R$ , we write  $H_R$  in place of  $H$ , so that  $H_R(t) = t + R(t)$ . Using the standard notation if  $R = (R_1, \dots, R_n)$ , we have

$$dR(t) = \begin{bmatrix} \frac{\partial R_1}{\partial t_1}(t) & \dots & \frac{\partial R_1}{\partial t_n}(t) \\ \vdots & \ddots & \vdots \\ \frac{\partial R_n}{\partial t_1}(t) & \dots & \frac{\partial R_n}{\partial t_n}(t) \end{bmatrix}.$$

Setting  $\widehat{Y}_j := (H_R)_* Y_j$ , a direct computation shows

$$\widehat{Y} = \frac{\partial}{\partial v} + \widehat{A}(v) \frac{\partial}{\partial v},$$

where

$$(6.7) \quad \widehat{A}(v) = \left( dR(t)^\top + A(t)(I + dR(t)^\top) \right) \Big|_{t=H_R^{-t}(v)}, \quad v \in H_R(B^n(4)).$$

Without loss of generality, we take  $\sigma \leq \frac{\gamma_1}{2}$ , and by taking  $\gamma_2 > 0$  sufficiently small (6.7) implies (f).

We wish to pick  $R$  so that

$$(6.8) \quad \sum_{j=1}^n \frac{\partial}{\partial v_j} \hat{A}_j(v) = 0, \quad v \in H_R(B^n(4)).$$

Define  $\Psi(A, R)(t) := (\Psi_1(A, R)(t), \dots, \Psi_n(A, R)(t))$  by

$$\begin{aligned} & \Psi_k(A, R)(t) \\ &:= \sum_{j=1}^n \frac{\partial}{\partial v_j} \left( dR(H_R^{-1}(v))^\top + A(H_R^{-1}(v)) \left( I + dR(H_R^{-1}(v))^\top \right) \right)_{j,k} \Big|_{v=H_R(t)}; \end{aligned}$$

where the subscript  $j, k$  denotes taking the  $(j, k)$  component of the matrix. In light of (6.7), (6.8) is equivalent to  $\Psi(A, R)(t) = 0$ ,  $t \in B^n(4)$ .

For any function  $K(t)$ , the chain rule shows

$$(6.9) \quad \frac{\partial}{\partial v_j} K(H_R^{-1}(v)) \Big|_{v=H_R(t)} = dK(t) (I + dR(t))^{-1} e_j,$$

where  $e_j$  denotes the  $j$ th standard basis element—the point is that the right-hand side of (6.9) is a function of  $dK(t)$  and  $dR(t)$ . Thus, using the notation of Appendix A.3, we have

$$\Psi(A, R)(t) = g(\mathcal{D}^1 A(t), \mathcal{D}^2 R(t))$$

for some smooth function  $g$  defined near the origin, with  $g(0, 0) = 0$ . Furthermore, it is easy to see that  $g(\mathcal{D}^1 A(t), \mathcal{D}^2 R(t))$  is quasilinear in  $R$  in the sense of (A.10).

We wish to solve for  $R$  in terms of  $A$  so that  $\Psi(A, R) = 0$ , provided  $\|A\|_{\mathcal{C}^0(B^n(5); \mathbb{M}^{n \times n})} \leq \gamma_2$ , where  $\gamma_2$  is a chosen small as in the statement of the lemma. To do this, we apply Proposition A.4; thus we need to make sure  $g(\mathcal{D}^1 A(t), \mathcal{D}^2 R(t))$  is elliptic in the sense of that proposition (where we are replacing  $B$  with  $R$  in the statement of that proposition). Define  $\mathcal{E}_2$  as in (A.11); we wish to show  $\mathcal{E}_2$  is elliptic. Note that

$$R \longmapsto \frac{d}{d\epsilon} \Big|_{\epsilon=0} \Psi(0, \epsilon R)$$

is a second order, constant coefficient, differential operator acting on  $R$  whose principal symbol is  $\mathcal{E}_2$ . Thus, we wish to show that this differential operator is elliptic. It suffices to compute this operator in the special case when  $R \in C^\infty$ .

Assuming  $R$  is  $C^\infty$ , we have

$$H_{\epsilon R}(t) = t + \epsilon R(t), \quad H_{\epsilon R}^{-1}(v) = v - \epsilon R(v) + O(\epsilon^2),$$

where  $O(\epsilon^2)$  denotes a term which is  $C^\infty$  in the variable  $t$  or  $v$ , and all of whose derivatives in this variable (of all orders  $\geq 0$ ) are  $O(\epsilon^2)$ . Thus,

$$\begin{aligned}\Psi_k(0, \epsilon R) &= \sum_{j=1}^n \frac{\partial}{\partial v_j} \left( \epsilon dR(H_{\epsilon R}^{-1}(v))^\top \right)_{j,k} \Big|_{v=H_{\epsilon R}(t)} \\ &= \sum_{j=1}^n \frac{\partial}{\partial v_j} \left( \epsilon dR(v)^\top \right)_{j,k} \Big|_{v=t+\epsilon R(t)} + O(\epsilon^2) \\ &= \sum_{j=1}^n \epsilon \frac{\partial}{\partial v_j} \frac{\partial R_k}{\partial v_j}(v) \Big|_{v=t+\epsilon R(t)} + O(\epsilon^2) \\ &= \sum_{j=1}^n \epsilon \frac{\partial^2}{\partial t_j^2} R_k(t) + O(\epsilon^2).\end{aligned}$$

Thus,

$$\frac{d}{d\epsilon} \Big|_{\epsilon=0} \Psi(0, \epsilon R) = \left( \sum_{j=1}^n \frac{\partial^2}{\partial t_j^2} R_1, \sum_{j=1}^n \frac{\partial^2}{\partial t_j^2} R_2, \dots, \sum_{j=1}^n \frac{\partial^2}{\partial t_j^2} R_n \right),$$

and we conclude  $g(\mathcal{D}^1 A(t), \mathcal{D}^2 R(t))$  is elliptic in the sense of Proposition A.4.

We apply Proposition A.4 with  $D = 4$ ,  $\eta = 3$ , and

$$N = \{R \in \mathcal{C}^{s_0+1}(B^n(4); \mathbb{R}^n) : \|R\|_{\mathcal{C}^{s_0+1}(B^n(4); \mathbb{R}^n)} < \sigma\}.$$

Thus, if  $\gamma_2 > 0$  is sufficiently small, and if  $\|A\|_{\mathcal{C}^{s_0}(B^n(5); \mathbb{M}^{n \times n})} \leq \gamma_2$ , we may solve for  $R = R(A) \in N$  such that  $\Psi(A, R) = 0$ ,  $R(0) = 0$ ,  $dR(0) = 0$ , and (c) and (d) hold. As we saw earlier,  $\Psi(A, R) = 0$  is equivalent to (e), and (a) and (f) have already been verified. This completes the proof.  $\square$

*Remark 6.10.* Throughout this paper, we fixed  $s_0 > 1$ . It would be nice if we could achieve the same results for  $s_0 > 0$ , however technical issues arise if we try to follow the methods of this paper with  $s_0 \in (0, 1]$ . This is particularly notable in the proof of Lemma 6.9. When  $s_0 > 1$ , the solutions we consider to the PDE which arises in that lemma are classical, however if  $s_0 \in (0, 1]$ , it seems likely one would have to consider some kind of generalized solution. A similar problem occurs in the proof of Proposition 6.6.

*Proof of Proposition 6.8.* Let  $\gamma_1 = \gamma_1(n) > 0$  be as in Proposition 6.6. We shrink  $\gamma_1 > 0$ , if necessary, to ensure that if  $\hat{A}$  is an  $n \times n$  matrix with components  $\hat{a}_j^k$  and  $|\hat{a}_j^k| \leq \gamma_1$ , then  $\|\hat{A}\|_{\mathbb{M}^{n \times n}} \leq \frac{1}{2}$ . We take  $\sigma_{n, s_0} > 0$  to be so small that if  $\|R\|_{\mathcal{C}^{s_0+1}(B^n(4); \mathbb{R}^n)} \leq \sigma_{n, s_0}$  we have

- If  $H(t) = t + R(t)$ , then  $B^n(2) \subseteq H(B^n(3))$ .
- $\det dH(t) \geq \frac{1}{2}$ ,  $\forall t \in B^n(3)$ .

Applying Lemma 6.9 with this choice of  $\gamma_1$  and with  $\sigma = \sigma_{n,s_0}$  yields  $\gamma_2$  and  $H$  as in that theorem. Since  $B^n(2) \subseteq H(B^n(3))$ , by the choice of  $\sigma_{n,s_0}$ , and in light of Lemma 6.9 (a), we may define  $\Phi_1 : B^n(2) \rightarrow B^n(3) \subseteq B^n(5)$  by  $\Phi_1(t) = H^{-1}(t)$ . (c), (d), and (e) follow from the corresponding properties of  $H$  described in Lemma 6.9.

Since  $\|H\|_{\mathcal{C}^{s+1}(B^n(3);\mathbb{R}^n)} \lesssim_{\{s\}} 1$  (by Lemma 6.9 (d)) and because  $\det dH(t) \geq \frac{1}{2}$ ,  $\forall t \in B^n(3)$  (by the choice of  $\sigma = \sigma_{n,s_0}$ ), we have  $\|\Phi_1\|_{\mathcal{C}^{s+1}(B^n(2);\mathbb{R}^n)} \lesssim_{\{s\}} 1$  (see Lemma 5.9), proving (b); the same proof gives (a). Moreover, if  $\widehat{Y}_j = \Phi_1^* Y_j = H_* Y_j$ , we have  $\|\widehat{Y}_j\|_{\mathcal{C}^s(B^n(2);\mathbb{R}^n)} \lesssim_{\{s\}} 1$ . Writing  $\widehat{Y} = \frac{\partial}{\partial t} + \widehat{A} \frac{\partial}{\partial t}$ , that  $\widehat{A}(0) = 0$  follows from (c) and the fact that  $A(0) = 0$ . That  $\sup_{u \in B^n(1)} \|\widehat{A}(u)\|_{\mathbb{M}^{n \times n}} \leq \frac{1}{2}$  follows from the choice of  $\gamma_1$  and Lemma 6.9 (f). This establishes (f).

All that remains to establish are the two (clearly equivalent) statements (g) and (h). For this, we use Proposition 6.6. Since  $\|\widehat{Y}_j\|_{\mathcal{C}^s(B^n(2);\mathbb{R}^n)} \lesssim_{\{s\}} 1$ , we have  $\|\widehat{a}_j^k\|_{\mathcal{C}^s(B^n(2))} \lesssim_{\{s\}} 1$ . Also, we have

$$[\widehat{Y}_j, \widehat{Y}_k] = \Phi_1^* [Y_j, Y_k] = \Phi_1^* \sum_{l=1}^n c_{j,k}^l Y_l = \sum_{l=1}^n \widehat{c}_{j,k}^l \widehat{Y}_l,$$

where  $\widehat{c}_{j,k}^l = c_{j,k}^l \circ \Phi_1$ . Using (b), Lemma 5.8, and the assumption  $\|c_{j,k}^l\|_{\mathcal{C}^s(B^n(5))} \lesssim_{\{s\}} 1$ , this implies  $\|\widehat{c}_{j,k}^l\|_{\mathcal{C}^s(B^n(2))} \lesssim_{\{s\}} 1$ . Finally, Lemma 6.9 (e) and (f) show that all of the hypotheses of Proposition 6.6 hold for  $\widehat{Y}_1, \dots, \widehat{Y}_n$ . Applying Proposition 6.6 yields (g) and (h), completing the proof.  $\square$

**6.6. Construction of  $\Phi_2$ .** In this section, we prove Proposition 6.3, and we take the same setting and notation as in that proposition; thus, we have vector fields  $Y_1, \dots, Y_n$  and functions  $\tilde{c}_{i,j}^k$  as in that proposition, and we have a notion of  $\{s\}$ -admissible constants given in Definition 6.1. Because of this definition of  $\{s\}$ -admissible constants, it suffices to assume  $s \geq s_0$  in all of Proposition 6.3. Thus, in this section we consider only  $s \geq s_0$ .

**LEMMA 6.11.** *Define, for  $\gamma \in (0, 1]$ ,  $\Psi_\gamma : B^n(\eta_1/\gamma) \rightarrow B^n(\eta_1)$  by  $\Psi_\gamma(t) = \gamma t$ . Let  $Y_j^\gamma := \gamma \Psi_\gamma^* Y_j$ . Then,  $Y_j^\gamma = \frac{\partial}{\partial t} + A_\gamma \frac{\partial}{\partial t}$  and  $[Y_j^\gamma, Y_k^\gamma] = \sum_{l=1}^n c_{j,k}^{l,\gamma} Y_l^\gamma$ , where for  $\gamma \in (0, \min\{\frac{\eta_1}{5}, 1\}]$ ,  $s \geq s_0$ ,*

$$(6.10) \quad \|A_\gamma\|_{\mathcal{C}^s(B^n(5);\mathbb{M}^{n \times n})} \lesssim_{\{s\}} \gamma, \quad \|c_{j,k}^{l,\gamma}\|_{\mathcal{C}^s(B^n(5))} \lesssim_{\{s\}} \gamma.$$

*Proof.* Since  $A_\gamma(t) = A(\gamma t)$  and  $A(0) = 0$ , that  $\|A_\gamma\|_{\mathcal{C}^s(B^n(5);\mathbb{M}^{n \times n})} \lesssim_{\{s\}} \gamma$  follows from Lemma 5.10 (this uses  $s \geq s_0 > 1$ ). Since  $c_{j,k}^{l,\gamma}(t) = \gamma c_{j,k}^l(\gamma t)$ ,  $\|c_{j,k}^{l,\gamma}\|_{\mathcal{C}^s(B^n(5))} \lesssim_{\{s\}} \gamma$  follows directly from the definitions (this uses  $\gamma \in (0, 1]$ ).  $\square$

*Proof of Proposition 6.3.* Let  $A_\gamma$ ,  $c_{j,k}^{l,\gamma}$ , and  $Y_j^\gamma$  be as in Lemma 6.11. Fix  $\gamma_2 = \gamma_2(n, s_0) > 0$  as in Proposition 6.8. Take  $\gamma \approx_{\{s_0\}} 1$  so small  $\gamma \leq \min\{\frac{\eta_1}{5}, 1\}$  and  $\|A_\gamma\|_{\mathcal{C}^{s_0}(B^n(5);\mathbb{M}^{n \times n})} \leq \gamma_2$  (this is clearly possible by (6.10)). With this choice of  $\gamma$ , we have  $\|c_{j,k}^{l,\gamma}\|_{\mathcal{C}^s(B^n(5))} \lesssim_{\{s\}} \gamma \leq 1$  and  $\|A_\gamma\|_{\mathcal{C}^s(B^n(5))} \lesssim_{\{s\}} \gamma \leq 1$ , for  $s \geq s_0$ , by (6.10).

In light of these remarks, Proposition 6.8 applies to  $Y_1^\gamma, \dots, Y_n^\gamma$  to yield a map  $\Phi_1 : B^n(1) \rightarrow B^n(5)$  as in that proposition (and constants which are  $\{s\}$ -admissible in the sense of that proposition are  $\{s\}$ -admissible in the sense of this section). Let  $\widehat{Y}_j^\gamma = \Phi_1^* Y_j^\gamma$ .

Set  $\Phi_2 := \Psi_\gamma \circ \Phi_1 : B^n(1) \rightarrow B^n(\eta_1)$ , and let  $\widehat{Y}_j = \Phi_2^* Y_j$ . Note that  $\widehat{Y}_j = K \widehat{Y}_j^\gamma$ , where  $K := \frac{1}{\gamma} \geq 1$  is an  $\{s_0\}$ -admissible constant. With this choice of  $K$  and  $\Phi_2$ , the proposition follows from the corresponding results for  $\Phi_1$  and  $\widehat{Y}_1^\gamma, \dots, \widehat{Y}_n^\gamma$  given in Proposition 6.8.  $\square$

**6.7. Qualitative results.** We now turn to the qualitative results; i.e., Theorems 2.4 and 2.7. These are simple consequences of Theorem 2.14. We begin with Theorem 2.4. For this we recall [SS18, Proposition 4.14].

LEMMA 6.12 (Proposition 4.14 of [SS18]). *Let  $X_1, \dots, X_q$  be  $C^1$  vector fields on a  $C^2$  manifold  $\mathfrak{M}$ .*

- $\forall x_0 \in \mathfrak{M}, \exists \eta > 0$ , such that  $X_1, \dots, X_q$  satisfy  $\mathcal{C}(x_0, \eta, \mathfrak{M})$ .
- Let  $K \subseteq \mathfrak{M}$  be a compact set. Then, there 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  $\forall r \in (0, \delta_0]$ ,

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

*Remark 6.13.* Lemma 6.12 shows that we always have  $\eta$  and  $\delta_0$  as in the assumptions of Theorem 2.14. Thus, if we wish to apply Theorem 2.14 to obtain a qualitative result, we do not need to verify the existence of  $\eta$  and  $\delta_0$ .

*Proof of Theorem 2.4.* (i)  $\Rightarrow$  (ii): First we prove the result with  $s < \infty$ . Let  $U$ ,  $V$ ,  $x_0$ , and  $\Phi$  be as in (i). Without loss of generality assume  $0 \in U$  and  $\Phi(0) = x_0$ . Reorder  $X_1, \dots, X_q$  so that  $X_1(x_0), \dots, X_n(x_0)$  are linearly independent and let  $Y_j = \Phi^* X_j$ , so that  $Y_j \in \mathcal{C}^{s+1}(U; \mathbb{R}^n)$ ,  $1 \leq j \leq q$ . Note that  $Y_1(0), \dots, Y_n(0)$  span the tangent space  $T_0 U$ . Let  $\eta > 0$  be so small  $B^n(2\eta) \subset U$  and  $Y_1, \dots, Y_n$  form a basis for the tangent space on  $B^n(2\eta)$ . It is immediate to verify, for  $1 \leq j, k \leq q$ , that

$$(6.11) \quad [Y_j, Y_k] = \sum_{l=1}^n \tilde{c}_{j,k}^l Y_l,$$

where  $\tilde{c}_{j,k}^l \in \mathcal{C}^s(B^n(\eta))$ . Because  $Y_1, \dots, Y_q$  span the tangent space at every point of  $B^n(2\eta)$  and  $Y_j \in \mathcal{C}^{s+1}(B^n(2\eta); \mathbb{R}^n)$ ,  $1 \leq j \leq q$ , Corollary 5.13 implies

$$(6.12) \quad \tilde{c}_{i,j}^k \in \mathcal{C}^s(B^n(\eta)) = \mathcal{C}_Y^s(B^n(\eta)).$$

Pushing (6.11) forward via  $\Phi$  shows  $[X_j, X_k] = \sum_{l=1}^n \tilde{c}_{j,k}^l X_l$ , with  $\tilde{c}_{j,k}^l = \tilde{c}_{j,k}^l \circ \Phi^{-1}$ . (2.3) and (6.12) combine to give  $\tilde{c}_{j,k}^l \in \mathcal{C}_X^s(\Phi(B^n(\eta)))$ .

Using that  $Y_1, \dots, Y_n$  span the tangent space at every point of  $B^n(2\eta)$  and that  $Y_j \in \mathcal{C}^{s+1}(U; \mathbb{R}^n)$ ,  $1 \leq j \leq q$ , for  $n+1 \leq j \leq q$ , we may write

$$(6.13) \quad Y_j = \sum_{k=1}^n \tilde{b}_j^k Y_k$$

where  $\tilde{b}_j^k \in \mathcal{C}^{s+1}(B^n(\eta))$ . By Corollary 5.13,  $\tilde{b}_j^k \in \mathcal{C}^{s+1}(B^n(\eta)) = \mathcal{C}_Y^{s+1}(B^n(\eta))$ , and by (2.3),  $b_j^k = \tilde{b}_j^k \circ \Phi^{-1} \in \mathcal{C}_X^{s+1}(\Phi(B^n(\eta)))$ . Pushing (6.13) forward via  $\Phi$ , we have  $X_j = \sum_{k=1}^n b_j^k X_k$  on  $\Phi(B^n(\eta))$  this completes the proof of (ii) with  $V$  replaced by  $\Phi(B^n(\eta))$ , when  $s < \infty$ .

If  $s = \infty$  note that in the above proof  $\eta$ ,  $\tilde{c}_{j,k}^l$ , and  $b_j^k$  can be chosen independent of  $s$ , thus when  $s = \infty$  the above proof applied to each  $s < \infty$  completes the proof of (ii).

(ii)  $\Rightarrow$  (iii): Suppose (ii) holds. We wish to show for  $1 \leq i, j \leq q$ ,

$$(6.14) \quad [X_i, X_j] = \sum_{k=1}^q c_{i,j}^k X_k, \quad c_{i,j}^k \in \mathcal{C}_X^s(V).$$

where  $s$  and  $V$  are as in (ii). For  $1 \leq i, j \leq n$ , (6.14) is contained in (ii). We prove the result for  $n+1 \leq i, j \leq q$ . The remaining cases ( $1 \leq i \leq n$  and  $n+1 \leq j \leq q$ , or  $n+1 \leq i \leq q$  and  $1 \leq j \leq n$ ) are similar and easier. We have

$$\begin{aligned} [X_i, X_j] &= \left[ \sum_{k_1=1}^n b_i^{k_1} X_{k_1}, \sum_{k_2=1}^n b_j^{k_2} X_{k_2} \right] \\ &= \sum_{k_1, k_2=1}^n \left( b_i^{k_1} (X_{k_1} b_j^{k_2}) X_{k_2} - b_j^{k_2} (X_{k_2} b_i^{k_1}) X_{k_1} + \sum_{l=1}^n b_i^{k_1} b_j^{k_2} \tilde{c}_{k_1, k_2}^l X_l \right). \end{aligned}$$

We are given  $b_j^k \in \mathcal{C}_X^{s+1}(V)$  and  $\tilde{c}_{k_1, k_2}^l \in \mathcal{C}_X^s(V)$ . It follows immediately from the definition of  $\mathcal{C}_X^{s+1}$  that  $X_l b_j^k \in \mathcal{C}_X^s(V)$ . From here, (6.14) follows from the fact that  $\mathcal{C}_X^s(V)$  is an algebra (see Lemma 5.7), completing the proof of (iii).

(iii)  $\Rightarrow$  (i): This is a consequence of Theorem 2.14. We make a few comments to this end. First of all, as discussed in Lemma 6.12 and Remark 6.13, there exist  $\eta$  and  $\delta_0$  as in the hypotheses of Theorem 2.14. Fix any  $s_0 \in (1, s] \setminus \{\infty\}$  and take  $\xi > 0$  so small  $B_X(x_0, \xi) \subseteq V$ . Take  $J_0$  as in Theorem 2.14 (with  $\zeta = 1$ ). We have,

directly from the definitions,

$$c_{i,j}^k \in \mathcal{C}_X^s(V) \subseteq \mathcal{C}_X^s(B_X(x_0, \xi)) \subseteq \mathcal{C}_{X_{J_0}}^s(B_{X_{J_0}}(x_0, \xi)) \subseteq \mathcal{C}_{X_{J_0}}^{s_0}(B_{X_{J_0}}(x_0, \xi)).$$

Thus, all of the hypotheses of Theorem 2.14 hold for this choice of  $s_0$ . This yields a map  $\Phi$  as in Theorem 2.14. This map satisfies the conclusions of (i), and this completes the proof.  $\square$

We now turn to Theorem 2.7. The uniqueness of the  $\mathcal{C}^{s+2}$  structure described in that theorem follows from the next lemma.

LEMMA 6.14. *Fix  $s \in (0, \infty]$ . Let  $M$  and  $N$  be two  $n$ -dimensional  $\mathcal{C}^{s+2}$  manifolds, and suppose  $X_1, \dots, X_q$  are  $\mathcal{C}^{s+1}$  vector fields on  $M$  which span the tangent space at every point, and  $Z_1, \dots, Z_q$  are  $\mathcal{C}^{s+1}$  vector fields on  $N$ . Let  $\Psi: M \rightarrow N$  be a  $C^2$  diffeomorphism such that  $\Psi_* X_j = Z_j$ . Then  $\Psi$  is a  $\mathcal{C}^{s+2}$  diffeomorphism.*

*Proof.* We first prove the result in the special case when  $M$  and  $N$  are open subsets of  $\mathbb{R}^n$ ; in this case we can identify the vector fields with  $\mathbb{R}^n$  valued functions, in the usual way. We use  $x$  to denote points in  $M \subseteq \mathbb{R}^n$  and  $y$  to denote a points in  $N \subseteq \mathbb{R}^n$ .

Fix a point  $x_0 \in M$ , we will show  $\Psi \in \mathcal{C}_{\text{loc}}^{s+1}$  on a neighborhood of  $x_0$ ; since  $x_0 \in M$  is arbitrary, this will complete the proof of the case when  $M$  and  $N$  are open subsets of  $\mathbb{R}^n$ . Reorder  $X_1, \dots, X_q$  so that  $X_1(x_0), \dots, X_n(x_0)$  are linearly independent; and reorder  $Z_1, \dots, Z_q$  in the same way to that we still have  $\Psi_* X_j = Z_j$ . Since  $X_1(x_0), \dots, X_n(x_0)$  form a basis of  $T_{x_0} M$ , we may pick an open neighborhood  $U$  of  $x_0$  so that  $X_1(x), \dots, X_n(x)$  form a basis for the tangent space at every  $x \in U$ .

Let  $\mathcal{X}(x) := (X_1(x) | \dots | X_n(x))$ ; i.e.,  $\mathcal{X}$  is the  $n \times n$  matrix whose columns are given by the vectors  $X_1, \dots, X_n$ . Similarly, let  $\mathcal{Z}(y) = (Z_1(y) | \dots | Z_n(y))$ . By hypothesis, we have  $\mathcal{X} \in \mathcal{C}_{\text{loc}}^{s+1}(M; \mathbb{M}^{n \times n})$  and  $\mathcal{Z} \in \mathcal{C}_{\text{loc}}^{s+1}(N; \mathbb{M}^{n \times n})$ . Since  $\Psi_* X_j = Z_j$ , we have the matrix equation

$$(6.15) \quad d\Psi(x) \mathcal{X}(x) = \mathcal{Z}(\Psi(x)), \quad x \in M.$$

Since  $X_1, \dots, X_n$  span the tangent space at every point of  $U$ , the matrix  $\mathcal{X}$  is invertible,  $\forall x \in U$ . It follows from Lemma 5.7 (by using the cofactor formula for  $\mathcal{X}(x)^{-1}$ ), that  $\mathcal{X}(\cdot)^{-1} \in \mathcal{C}_{\text{loc}}^{s+1}(U; \mathbb{M}^{n \times n})$ . From (6.15), we obtain

$$(6.16) \quad d\Psi(x) = \mathcal{Z}(\Psi(x)) \mathcal{X}(x)^{-1}, \quad x \in U.$$

Suppose  $\Psi \in \mathcal{C}_{\text{loc}}^{s'+2}(U; \mathbb{R}^n)$ , for some  $s' \geq 0$ . We will show

$$\Psi \in \mathcal{C}_{\text{loc}}^{\min\{s'+3, s+2\}}(U; \mathbb{R}^n);$$

and then it will follow by iteration that  $\Psi \in \mathcal{C}_{\text{loc}}^{s+2}(U; \mathbb{R}^n)$ , as desired. This will complete the proof since  $C_{\text{loc}}^2(U; \mathbb{R}^n) \subset \mathcal{C}_{\text{loc}}^2(U; \mathbb{R}^n)$ . Since  $\Psi \in \mathcal{C}_{\text{loc}}^{s'+2}(U; \mathbb{R}^n)$  and  $\mathcal{Z} \in \mathcal{C}_{\text{loc}}^{s+1}(N; \mathbb{M}^{n \times n})$ , it follows from Lemma 5.15 that

$$\mathcal{Z} \circ \Psi \in \mathcal{C}_{\text{loc}}^{\min\{s'+2, s+1\}}(U; \mathbb{M}^{n \times n}).$$

Since we have already shown  $\mathcal{X}(\cdot)^{-1} \in \mathcal{C}_{\text{loc}}^{s+1}(U; \mathbb{M}^{n \times n})$ , it follows from Lemma 5.7 and (6.16) that

$$d\Psi(x) = \mathcal{Z}(\Psi(x))\mathcal{X}(x)^{-1} \in \mathcal{C}_{\text{loc}}^{\min\{s'+2, s+1\}}(U; \mathbb{M}^{n \times n}).$$

Since we also have  $\Psi \in \mathcal{C}_{\text{loc}}^{s'+2}(U; \mathbb{R}^n)$ , it follows that  $\Psi \in \mathcal{C}_{\text{loc}}^{\min\{s'+3, s+2\}}(U; \mathbb{R}^n)$ , as desired. This completes the proof in the case when  $M$  and  $N$  are open subsets of  $\mathbb{R}^n$ .

We now turn to the general case, where  $M$  and  $N$  are  $\mathcal{C}^{s+2}$  manifolds of dimension  $n$ , and  $X_1, \dots, X_q$ ,  $Z_1, \dots, Z_q$ , and  $\Psi$  are as in the statement of the lemma. Since  $M$  and  $N$  are  $\mathcal{C}^{s+2}$  manifolds they have associated  $\mathcal{C}^{s+2}$  atlases  $\{(\phi_\alpha, V_\alpha)\}$  and  $\{(\psi_\beta, W_\beta)\}$ , respectively. We wish to show,  $\forall \alpha, \beta$ ,

$$\Psi_{\alpha, \beta} := \psi_\beta \circ \Psi \circ \phi_\alpha^{-1} : \phi_\alpha(V_\alpha \cap \Psi^{-1}(W_\beta)) \longrightarrow \psi_\beta(\Psi(V_\alpha) \cap W_\beta)$$

is a  $\mathcal{C}^{s+2}$  diffeomorphism, and this will complete the proof.

By hypothesis, we have

$$(\Psi_{\alpha, \beta})_*((\phi_\alpha)_*X_j) = (\psi_\beta)_*Z_j.$$

Since  $(\phi_\alpha)_*X_1, \dots, (\phi_\alpha)_*X_q$  and  $(\psi_\beta)_*Z_1, \dots, (\psi_\beta)_*Z_q$  are  $\mathcal{C}^{s+1}$  vector fields, by hypothesis, and  $(\phi_\alpha)_*X_1, \dots, (\phi_\alpha)_*X_q$  span the tangent space at every point of  $\phi_\alpha(V_\alpha)$ , it follows from the above case (when  $M$  and  $N$  are open subsets of  $\mathbb{R}^n$ ), that  $\Psi_{\alpha, \beta}$  is a  $\mathcal{C}^{s+2}$  diffeomorphism. This completes the proof.  $\square$

*Proof of Theorem 2.7.* (ii) $\Rightarrow$ (i): Under the condition (ii), for each  $x \in M$ , there exist open sets  $U_x \subseteq \mathbb{R}^n$ ,  $V_x \subseteq M$ , and a  $C^2$  diffeomorphism  $\Phi_x : U_x \rightarrow V_x$  such that if  $Y_j^x = \Phi_x^*X_j$ , then  $Y_j^x \in \mathcal{C}^{s+1}(U_x; \mathbb{R}^n)$ . We wish to show that the collection  $\{(\Phi_x^{-1}, V_x) : x \in M\}$  forms a  $\mathcal{C}^{s+2}$  atlas on  $M$ ; once that is shown, (i) will follow since the  $X_j$  will be  $\mathcal{C}^{s+1}$  with respect to this atlas by definition, and this atlas is clearly compatible with the  $C^2$  structure on  $M$ . Hence, we need only verify that the transition functions are  $\mathcal{C}_{\text{loc}}^{s+2}$ . Take  $x_1, x_2 \in M$  such that  $V_{x_1} \cap V_{x_2} \neq \emptyset$ . Set  $\Psi = \Phi_{x_2}^{-1} \circ \Phi_{x_1} : U_{x_1} \cap \Phi_{x_1}^{-1}(V_{x_2}) \rightarrow U_{x_2} \cap \Phi_{x_2}^{-1}(V_{x_1})$ . We wish to show  $\Psi$  is a  $\mathcal{C}^{s+2}$  diffeomorphism. We already know  $\Psi$  is a  $C^2$  diffeomorphism and  $\Psi_*Y_j^{x_1} = Y_j^{x_2}$ . That  $\Psi$  is a  $\mathcal{C}^{s+2}$  diffeomorphism now follows from Lemma 6.14, completing the proof of (i).

(i) $\Rightarrow$ (iii): Suppose (i) holds. Using a simple partition of unity argument, we may write  $[X_j, X_k] = \sum_{l=1}^q c_{j,k}^l X_l$ , where  $c_{j,k}^l : M \rightarrow \mathbb{R}$  and are  $\mathcal{C}_{\text{loc}}^s$  maps.

We wish to show  $\forall x_0 \in M$ ,  $\exists V \subseteq M$  open with  $x_0 \in V$  and  $c_{j,k}^l|_V \in \mathcal{C}_X^s(V)$ . Fix  $x_0 \in M$ , and let  $W \subseteq M$  be a neighborhood of  $x_0$  such that there is a  $\mathcal{C}^{s+2}$  diffeomorphism  $\Phi : B^n(1) \rightarrow W$  with  $\Phi(0) = x_0$ . Set  $Y_j = \Phi^*X_j$ , so that  $Y_j \in \mathcal{C}^{s+1}(B^n(3/4); \mathbb{R}^n)$  and  $Y_1, \dots, Y_q$  span the tangent space at every point of  $B^n(1)$ . Also we have  $c_{j,k}^l \circ \Phi \in \mathcal{C}^s(B^n(3/4))$ . Corollary 5.13 shows  $c_{j,k}^l \circ \Phi \in \mathcal{C}^s(B^n(1/2)) = \mathcal{C}_Y^s(B^n(1/2))$  and (2.3) shows  $c_{j,k}^l \in \mathcal{C}_X^s(\Phi(B^n(1/2)))$ . This proves (ii) with  $V = \Phi(B^n(1/2))$ .

(iii)  $\Rightarrow$  (ii): This is obvious.

Finally, as mentioned before, the uniqueness of the  $\mathcal{C}^{s+2}$  manifold structure, as described in the theorem, is an immediate consequence of Lemma 6.14.  $\square$

**7. Hölder spaces.** Let  $\Omega \subset \mathbb{R}^n$  be a bounded, Lipschitz domain. It is easy to see that for  $m \in \mathbb{N}$ ,  $s \in [0, 1]$ ,  $m + s > 0$  we have the containment  $C^{m,s}(\Omega) \subseteq \mathcal{C}^{m+s}(\Omega)$ . For  $m \in \mathbb{N}$ ,  $s \in (0, 1)$ , we also have the reverse containment  $\mathcal{C}^{m+s}(\Omega) \subseteq C^{m,s}(\Omega)$ ; this follows easily from [Tri06, Theorem 1.118 (i)].

When we move to the corresponding spaces with respect to  $C^1$  vector fields  $X_1, \dots, X_q$  on a  $C^2$  manifold  $M$ , we have similar results. For any  $m \in \mathbb{N}$ ,  $s \in [0, 1]$  with  $m + s > 0$ , we have  $C_X^{m,s}(M) \subseteq \mathcal{C}_X^{m+s}(M)$ ; see [SS18, Lemma 8.1]. The reverse containment for  $m \in \mathbb{N}$  and  $s \in (0, 1)$  is a bit more difficult and requires appropriate hypotheses on the vector fields. We state a quantitative local version of this in the next proposition.

**PROPOSITION 7.1.** *We take all the same assumptions and notation as in Theorem 2.14, and let  $\Phi$  be as in that theorem (and  $\{s\}$ -admissible constants as in Definition 2.13). Then, for  $m \in \mathbb{N}$ ,  $s \in (0, 1)$ , and for any function  $f \in C(\Phi(B^n(1)))$ ,*

$$(7.1) \quad \|f\|_{C_X^{m,s}(\Phi(B^n(1)))} \approx_{\{m+s-2\}} \|f\|_{\mathcal{C}_X^{m+s}(\Phi(B^n(1)))}.$$

*Proof.* We use Lemma 5.6; in particular, for  $g \in C(B^n(1))$ ,  $m \in \mathbb{N}$ ,  $s \in (0, 1)$ ,

$$(7.2) \quad \|g\|_{C^{m,s}(B^n(1))} \approx \|g\|_{\mathcal{C}^{m+s}(B^n(1))},$$

where the implicit constants depend only on  $m + s$  and  $n$ . Let  $Y_j = \Phi^*X_j$ , and let  $A$  be as in Theorem 2.14. Letting  $Y_j = \Phi^*X_j$ , Theorem 2.14 (j) shows

$$\|Y_j\|_{\mathcal{C}^{m+s-1}(B^n(1); \mathbb{R}^n)} \lesssim_{\{m+s-2\}} 1,$$

and therefore by (7.2),  $\|Y_j\|_{C^{m-1,s}(B^n(1); \mathbb{R}^n)} \lesssim_{\{m+s-2\}} 1$ . Here, we are using the convention in Remark 5.11 to define  $C^{-1}$  and  $\mathcal{C}^{-1}$  when  $s - 1 \leq 0$ . Similarly, we have  $\|A\|_{\mathcal{C}^{m+s-1}(B^n(1); \mathbb{M}^{n \times n})}, \|A\|_{C^{m-1,s}(B^n(1); \mathbb{R}^n)} \lesssim_{\{m+s-2\}} 1$ .

Since  $Y_{J_0} = K(I + A)\nabla$  we have  $\nabla = K^{-1}(I + A)^{-1}Y_{J_0}$ . Thus, we may write  $\nabla$  as a linear combination of  $Y_1, \dots, Y_n$ , with coefficients whose  $C^{m-1,s}(B^n(1))$  and  $\mathcal{C}^{m+s-1}(B^n(1))$  norms are  $\lesssim_{\{m+s-2\}} 1$ .

With all of the above remarks, Proposition 5.12 shows for any  $g \in B^n(1)$ ,

$$\begin{aligned}\|g\|_{C^{m,s}(B^n(1))} &\approx_{\{m+s-2\}} \|g\|_{C_Y^{m,s}(B^n(1))}, \\ \|g\|_{\mathcal{C}^{m+s}(B^n(1))} &\approx_{\{m+s-2\}} \|g\|_{\mathcal{C}_Y^{m+s}(B^n(1))}.\end{aligned}$$

Combining this with (7.2), we have

$$(7.3) \quad \|g\|_{C_Y^{m,s}(B^n(1))} \approx_{\{m+s-2\}} \|g\|_{\mathcal{C}_Y^{m+s}(B^n(1))}.$$

(2.3) shows

$$\begin{aligned}\|f \circ \Phi\|_{C_Y^{m,s}(B^n(1))} &= \|f\|_{C_X^{m,s}(\Phi(B^n(1)))}, \\ \|f \circ \Phi\|_{\mathcal{C}_Y^{m+s}(B^n(1))} &= \|f\|_{\mathcal{C}_X^{m+s}(\Phi(B^n(1)))}.\end{aligned}$$

Combining this and (7.3) with  $g = f \circ \Phi$  yields (7.1) and completes the proof.  $\square$

Similarly, we may create Hölder versions of Theorems 2.4 and 2.7. We state these results here. We take the same setup as in Theorems 2.4 and 2.7.

**COROLLARY 7.2.** (The local result) *For  $m \in \mathbb{N}$ ,  $m \geq 1$  and  $s \in (0, 1)$  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  $\Phi^* X_1, \dots, \Phi^* X_q \in C^{m+1,s}(U; \mathbb{R}^n)$ .*

(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  $\hat{c}_{i,j}^k \in C_X^{m,s}(V)$ .
- For  $n+1 \leq j \leq q$ ,  $X_j = \sum_{k=1}^n b_j^k X_k$ , where  $b_j^k \in C_X^{m+1,s}(V)$ .

(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  $c_{i,j}^k \in C_X^{m,s}(V)$ .*

*Proof.* (i)  $\Rightarrow$  (ii)  $\Rightarrow$  (iii) has a nearly identical proof to the corresponding results in Theorem 2.4, and we leave the details to the reader. Assume (iii) holds. Then, since  $C_X^{m,s}(V) \subseteq \mathcal{C}_X^{m+s}(V)$  (by [SS18, Lemma 8.1]) we have that Theorem 2.4 (iii) holds (with  $s$  replaced by  $m+s$ ). Therefore, Theorem 2.4 (i) holds (again, with  $s$  replaced by  $m+s$ ); we may shrink  $U$  in Theorem 2.4 (i) so that it is a Euclidean ball. Letting  $\Phi$  be as in Theorem 2.4 (i), we have  $\Phi^* X_1, \dots, \Phi^* X_q \in \mathcal{C}^{m+s+1}(U; \mathbb{R}^n)$ . Since  $U$  is a ball, Lemma 5.6 shows  $\mathcal{C}^{m+s+1}(U; \mathbb{R}^n) = C^{m+1,s}(U; \mathbb{R}^n)$  (this is the point where we use  $s \neq 0, 1$ ). (i) follows, completing the proof.  $\square$

**Remark 7.3.** The only place  $m \geq 1$ ,  $s \neq 0, 1$  was used in Corollary 7.2 was (iii)  $\Rightarrow$  (i). The implications (i)  $\Rightarrow$  (ii)  $\Rightarrow$  (iii) hold for  $m \in \mathbb{N}$ ,  $s \in [0, 1]$  with the same proof. We do not know whether (iii)  $\Rightarrow$  (i) holds for  $m = 0$  or  $s = 0, 1$ .

**COROLLARY 7.4.** (The global result) *For  $m \in \mathbb{N}$ ,  $m \geq 1$  and  $s \in (0, 1)$ , the following three conditions are equivalent.*

- (i) *There exists a  $C^{m+2,s}$  atlas on  $M$ , compatible with its  $C^2$  structure, such that  $X_1, \dots, X_q$  are  $C^{m+1,s}$  with respect to this atlas.*
- (ii) *For each  $x_0 \in M$ , any of the three equivalent conditions from Corollary 7.2 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  $c_{i,j}^k|_V \in C_X^{m,s}(V)$ ,  $1 \leq i, j, k \leq q$ .

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

*Proof.* With Corollary 7.2 in hand, the proof is nearly identical to the proof of Theorem 2.7 and we leave the details to the reader.  $\square$

**Appendix A. Elliptic PDEs.** We require quantitative versions of some standard results from elliptic PDEs. The proofs of these results are well known, and the quantitative versions follow by keeping track of constants in the proofs. We make no effort to present the results or proofs in greatest generality, and only present what is needed for this paper.

**A.1. Regularity of linear elliptic equations.** Let  $\mathcal{E}$  be a constant coefficient partial differential operator of order  $M$ ,

$$\mathcal{E} : C^\infty(\mathbb{R}^n; \mathbb{C}^{m_1}) \longrightarrow C^\infty(\mathbb{R}^n; \mathbb{C}^{m_2}),$$

where  $m_2 \geq m_1$ . We may think of  $\mathcal{E}$  as a  $m_2 \times m_1$  matrix of constant coefficient partial differential operators of order  $\leq M$ .

Fix  $D \in (0, \infty)$ . Let  $\mathcal{L} = \sum_{|\alpha| \leq M} c_\alpha(x) \partial_x^\alpha$  where  $c_\alpha : B^n(D) \rightarrow \mathbb{M}^{m_2 \times m_1}(\mathbb{C})$ . For  $u : B^n(D) \rightarrow \mathbb{C}^{m_1}$  and  $g : B^n(D) \rightarrow \mathbb{C}^{m_2}$  we consider the equation

$$(A.1) \quad (\mathcal{E} + \mathcal{L})u = g.$$

**PROPOSITION A.1.** *Suppose  $\mathcal{E}$  is elliptic, and fix  $\epsilon_0 > 0$ . There exists  $\gamma = \gamma(\mathcal{E}) > 0$  such that if  $u$  and  $g$  satisfy (A.1) and  $\|c_\alpha\|_{L^\infty(B^n(D); \mathbb{M}^{m_2 \times m_1})} \leq \gamma$ ,  $\forall \alpha$ , then the following holds for all  $s > s_0 > 0$ ,  $\eta \in (0, D)$ ,*

$$(A.2) \quad \begin{aligned} u &\in \mathcal{C}^{s_0+M}(B^n(D); \mathbb{C}^{m_1}), \quad g \in \mathcal{C}^s(B^n(D); \mathbb{C}^{m_2}), \\ c_\alpha &\in \mathcal{C}^{s+\epsilon_0}(B^n(D); \mathbb{M}^{m_2 \times m_1}) \implies u \in \mathcal{C}^{s+M}(B^n(\eta); \mathbb{C}^{m_1}). \end{aligned}$$

Moreover we have

$$(A.3) \quad \|u\|_{\mathcal{C}^{s+M}(B^n(\eta); \mathbb{C}^{m_1})} \leq C \left( \|g\|_{\mathcal{C}^s(B^n(D); \mathbb{C}^{m_2})} + \|u\|_{\mathcal{C}^{s_0+M}(B^n(D); \mathbb{C}^{m_1})} \right),$$

where  $C$  can be chosen to depend only on  $s_0$ ,  $s$ ,  $\mathcal{E}$ ,  $D$ ,  $\eta$ ,  $\epsilon_0$ , and upper bounds for  $\|c_\alpha\|_{\mathcal{C}^{s+\epsilon_0}(B^n(D); \mathbb{M}^{m_2 \times m_2})}$ ,  $\|u\|_{\mathcal{C}^{s_0+M}(B^n(D))}$ , and  $\|g\|_{\mathcal{C}^s(B^n(D); \mathbb{C}^{m_2})}$ .

*Proof sketch.* We sketch a proof of (A.2) using theory from [Tay11]. There are many proofs of this result which are well known to experts. We use the theory from [Tay11] because that reference uses Zygmund spaces, while many other references only state results for Hölder spaces with non-integer exponents (even though many of these proofs can be generalized to Zygmund spaces). The quantitative estimate, (A.3), follows by keeping track of constants in this proof. For Zygmund spaces, [Tay11] uses the notation  $C_*^s$  instead of  $\mathcal{C}^s(\mathbb{R}^n)$ —for this proof, we use this notation to help the reader make the connection with the results in that book.

Note that if  $\gamma = \gamma(\mathcal{E}) > 0$  is sufficiently small,  $\mathcal{E} + \mathcal{L}$  is uniformly elliptic on  $B^n(D)$ . Let  $u \in \mathcal{C}^{s_0+M}(B^n(D); \mathbb{C}^{m_1})$ ,  $g \in \mathcal{C}^s(B^n(D); \mathbb{C}^{m_2})$ , and  $c_\alpha \in \mathcal{C}^{s+\epsilon_0}(B^n(D); \mathbb{M}^{m_2 \times m_1})$  satisfying (A.1). Fix  $\eta \in (0, D)$  and take  $\phi_1, \phi_2, \phi_3 \in C_0^\infty(B^n(D))$  such that  $\phi_j \equiv 1$  on a neighborhood of the support of  $\phi_{j-1}$  and  $\phi_1 \equiv 1$  on a neighborhood of the closure of  $B^n(\eta)$ . Since  $(\mathcal{E} + \mathcal{L})u = g$ , we have

$$(A.4) \quad \phi_2(\mathcal{E} + \mathcal{L})\phi_3 u = \phi_2 g.$$

Using the notation of Chapter 13, Section 9 of [Tay11], we have  $\phi_2(\mathcal{E} + \mathcal{L}) = a(x, D)$  where  $a(x, \xi) \in C_*^{s+\epsilon_0} S_{1,0}^M$ .

Set  $\delta = \min\left\{\frac{\epsilon_0}{s+\epsilon_0}, \frac{s-s_0}{s+\epsilon_0}\right\}$  so that  $\delta \in (0, 1)$ . By Proposition 9.9 of Chapter 13 of [Tay11],

$$a(x, \xi) = a^\sharp(x, \xi) + a^\flat(x, \xi), \quad a^\sharp \in S_{1,\delta}^M, \quad a^\flat \in C_*^{s+\epsilon_0} S_{1,\delta}^{M-(s+\epsilon_0)\delta}.$$

Note that since  $\mathcal{E} + \mathcal{L}$  is elliptic on  $B^n(D)$ ,  $a$  is elliptic on a neighborhood of the support of  $\phi_1$ , and the same is therefore true of  $a^\sharp$ .

Rewriting (A.4) we have

$$(A.5) \quad a^\sharp(x, D)\phi_3 u = \phi_2 g - a^\flat(x, D)\phi_3 u.$$

Since  $\phi_3 u \in C_*^{s_0+M}$ , by assumption, Proposition 9.10 of Chapter 13 of [Tay11] implies  $a^\flat(x, D)\phi_3 u \in C_*^{s_0+\min\{\epsilon_0, s-s_0\}}$ . Combining this with  $\phi_2 g \in C_*^s$  we have  $a^\sharp(x, D)\phi_3 u \in C_*^{s_0+\min\{\epsilon_0, s-s_0\}}$ .

Since  $a^\sharp$  is elliptic on a neighborhood of the support of  $\phi_1$ , we conclude  $\phi_1 u \in C_*^{s_0+M+\min\{\epsilon_0, s-s_0\}}$ , and therefore  $u \in \mathcal{C}^{s_0+M+\min\{\epsilon_0, s-s_0\}}(B^n(\eta); \mathbb{C}^{m_1})$ . (A.2) follows by iterating this result.  $\square$

*Remark A.2.* In [Tay11] a different (but equivalent) norm is used in the definition of  $\mathcal{C}^s(B^n(\eta))$  (see Remark 5.1). The constants in this equivalence depend on  $s$ ,  $n$ , and  $\eta$ . This does not create a problem in Proposition A.1 since  $C$  is allowed to depend on  $\mathcal{E}$  (and therefore on  $n$ ),  $s$ ,  $s_0$ ,  $\eta$ , and  $D$ .

**A.2. Regularity for a nonlinear elliptic equation.** Let  $\mathcal{E}$  be a constant coefficient, first order, partial differential operator,

$$\mathcal{E} : C^\infty(\mathbb{R}^n; \mathbb{C}^{m_1}) \longrightarrow C^\infty(\mathbb{R}^n; \mathbb{C}^{m_2}),$$

where  $m_2 \geq m_1$ . We may think of  $\mathcal{E}$  as a  $m_2 \times m_1$  matrix of constant coefficient partial differential operators of order  $\leq 1$ .

Let  $\Gamma : \mathbb{C}^{m_1} \times \mathbb{C}^{nm_1} \rightarrow \mathbb{C}^{m_2}$  be a bilinear map. Fix  $D > 0$ , we consider the equation, for  $b : B^n(D) \rightarrow \mathbb{C}^{m_1}$ ,  $c : B^n(D) \rightarrow \mathbb{C}^{m_2}$ ,

$$(A.6) \quad \mathcal{E}b + \Gamma(b, \nabla b) = c.$$

**PROPOSITION A.3.** *Suppose  $\mathcal{E}$  is elliptic. Then, there exists  $\gamma = \gamma(\mathcal{E}, \Gamma) > 0$  such that if  $b$  and  $c$  satisfy (A.6), and if for some  $s_1, s_2 > 0$  we have  $c \in \mathcal{C}^{s_2}(B^n(D); \mathbb{C}^{m_2})$ ,  $b \in \mathcal{C}^{s_1+1}(B^n(D); \mathbb{C}^{m_1})$ , with  $\|b\|_{L^\infty(B^n(D); \mathbb{C}^{m_1})} \leq \gamma$ , then for all  $\eta \in (0, D)$ ,  $b \in \mathcal{C}^{s_2+1}(B^N(\eta); \mathbb{C}^{m_1})$ . Moreover,*

$$\|b\|_{\mathcal{C}^{s_2+1}(B^n(\eta); \mathbb{C}^{m_1})} \leq C \left( \|b\|_{\mathcal{C}^{s_1+1}(B^n(D); \mathbb{C}^{m_1})} + \|c\|_{\mathcal{C}^{s_2}(B^n(D); \mathbb{C}^{m_2})} \right),$$

where  $C$  can be chosen to depend only on  $s_1, s_2, D, \eta, \mathcal{E}, \Gamma$ , and upper bounds for  $\|b\|_{\mathcal{C}^{s_1+1}(B^n(D); \mathbb{C}^{m_1})}$  and  $\|c\|_{\mathcal{C}^{s_2}(B^n(D); \mathbb{C}^{m_2})}$ .

*Proof.* We will show, under the hypotheses of the proposition, that there exists  $\gamma = \gamma(\mathcal{E}, \Gamma) > 0$  such that if  $b$  and  $c$  are as in the proposition, we have for  $\eta \in (0, D)$ ,

$$(A.7) \quad b \in \mathcal{C}^{\min\{s_1+3/2, s_2+1\}}(B^n(\eta); \mathbb{C}^{m_1}),$$

and

$$(A.8) \quad \|b\|_{\mathcal{C}^{\min\{s_1+3/2, s_2+1\}}(B^n(\eta); \mathbb{C}^{m_1})} \leq C \left( \|b\|_{\mathcal{C}^{s_1+1}(B^n(D); \mathbb{C}^{m_1})} + \|c\|_{\mathcal{C}^{s_2}(B^n(D); \mathbb{C}^{m_2})} \right),$$

where  $C$  is as in the statement of the proposition. The result then follows by iteration.

We use Proposition A.1 with  $M = 1$ ,  $\epsilon_0 = \frac{1}{2}$ ,  $s_0 = s_1$ , and  $s = \min\{s_2, s_1 + \frac{1}{2}\}$  applied to (A.6). With these choices, if  $\gamma = \gamma(\mathcal{E}, \Gamma) > 0$  is sufficiently small, Proposition A.1 applies to prove (A.7) and (A.8), completing the proof.  $\square$

**A.3. Existence for a nonlinear elliptic equation.** Fix  $D > 0$ ,  $m_1, m_2 \in \mathbb{N}$ . For functions  $A : B^n(D) \rightarrow \mathbb{R}^{m_1}$  and  $B : B^n(D) \rightarrow \mathbb{R}^{m_2}$  write

$$\mathcal{D}^1 A = (\partial_x^\alpha A)_{|\alpha| \leq 1}, \quad \mathcal{D}^2 B = (\partial_x^\alpha B)_{|\alpha| \leq 2}, \quad \mathcal{D}_2 B = (\partial_x^\alpha B)_{|\alpha|=2},$$

so that, for example,  $\mathcal{D}^2 B$  is the vector of all partial derivatives of  $B$  up to order 2, and  $\mathcal{D}_2 B$  is the vector of all partial derivatives of  $B$  of order exactly 2.

Fix a  $C^\infty$  function  $g$ . We wish to consider the equation

$$(A.9) \quad g(\mathcal{D}^1 A(x), \mathcal{D}^2 B(x)) = 0.$$

Here  $g$  is  $C^\infty$  and defined on a neighborhood of the origin, takes values in  $\mathbb{R}^{m_2}$ , and satisfies  $g(0, 0) = 0$ . Our goal is to give conditions on  $g$  so that given  $A$  (sufficiently small), we can find  $B = B(A)$  so that (A.9) holds; we further wish to understand the regularity properties of  $B$  in a quantitative way.

Though it is not necessary for the results that follow, we assume (A.9) is quasi-linear in  $B$ , which is sufficient for our purposes and simplifies the proof. That is, we assume

$$(A.10) \quad g(\mathcal{D}^1 A(x), \mathcal{D}^2 B(x)) = g_1(A(x), \mathcal{D}^1 B(x)) \mathcal{D}_2 B(x) + g_2(\mathcal{D}^1 A(x), \mathcal{D}^1 B(x)),$$

where  $g_1$  and  $g_2$  are smooth on a neighborhood of the origin,  $g_1$  takes values in matrices of an appropriate size, and  $g_2(0, 0) = 0$ .

Finally, let  $\mathcal{E}_2$  denote the second order partial differential operator

$$(A.11) \quad \mathcal{E}_2 B := g_1(0, 0) \mathcal{D}_2 B,$$

so that  $\mathcal{E}_2$  is an  $m_2 \times m_2$  matrix of constant, real coefficient partial differential operators of order  $\leq 2$ .

**PROPOSITION A.4.** *Suppose  $\mathcal{E}_2$  is elliptic. Fix  $s_0 > 0$  and a neighborhood  $N \subseteq \mathcal{C}^{2+s_0}(B^n(D); \mathbb{R}^{m_2})$  of 0. Then, there exists a neighborhood  $W \subseteq \mathcal{C}^{1+s_0}(B^n(D); \mathbb{R}^{m_2})$  of 0 and a map*

$$\mathcal{B} : W \longrightarrow N$$

such that

$$(A.12) \quad g(\mathcal{D}^1 A(x), \mathcal{D}^2 \mathcal{B}(A)(x)) = 0, \quad x \in B^n(D), A \in W.$$

This map satisfies  $\mathcal{D}^1 \mathcal{B}(A)(0) = 0$ ,  $\forall A \in W$ , and

$$(A.13) \quad \|\mathcal{B}(A)\|_{\mathcal{C}^{2+s_0}(B^n(D); \mathbb{R}^{m_2})} \leq C \|A\|_{\mathcal{C}^{1+s_0}(B^n(D); \mathbb{R}^{m_1})},$$

where  $C$  does not depend on  $A \in W$ . Finally, for  $\eta \in (0, D)$ , let  $R_\eta$  denote the restriction map  $R_\eta : f \mapsto f|_{B^n(\eta)}$ . Then, for  $s \geq s_0$ ,  $\eta \in (0, D)$ ,

$$(A.14) \quad R_\eta \circ \mathcal{B} : \mathcal{C}^{1+s}(B^n(D); \mathbb{R}^{m_1}) \cap W \rightarrow \mathcal{C}^{2+s}(B^n(\eta); \mathbb{R}^{m_2}),$$

and

$$(A.15) \quad \|R_\eta \circ \mathcal{B}(A)\|_{\mathcal{C}^{2+s}(B^n(\eta); \mathbb{R}^{m_2})} \leq C_{s, \eta},$$

where  $C_{s,\eta}$  can be chosen to depend on an upper bound for  $\|A\|_{\mathcal{C}^{1+s}(B^n(D);\mathbb{R}^{m_1})}$  and does not depend on  $A \in W$  in any other way. It can depend on any of the other ingredients in the problem.

The rest of this section is devoted to a sketch of a proof of Proposition A.4. The proof is a standard application of the Inverse Function Theorem combined with Proposition A.1; we include the proof as it gives the required quantitative estimates, which are essential for our purposes.

By expanding  $g$  into a Taylor series, we have

$$g(\mathcal{D}^1 A, \mathcal{D}^2 B) = \mathcal{A}A + \mathcal{E}B + q(\mathcal{D}^1 A, \mathcal{D}^2 B),$$

where  $\mathcal{A}$  is a first order linear differential operator with constant coefficients,  $\mathcal{E}$  is a second order linear differential operator with constant coefficients whose principal symbol is  $\mathcal{E}_2$ , and  $q$  is smooth and vanishes to second order at  $(0,0)$ .

Since  $\mathcal{E}$  is elliptic (because  $\mathcal{E}_2$  is), it is a standard fact that  $\mathcal{E}$  has a continuous right inverse

$$\mathcal{P} : \mathcal{C}^{s_0}(B^n(D);\mathbb{R}^{m_2}) \longrightarrow \mathcal{C}^{2+s_0}(B^n(D);\mathbb{R}^{m_2}),$$

where  $\mathcal{E}\mathcal{P} = I$  and for all  $|\alpha| \leq 1$ ,  $\partial_x^\alpha \mathcal{P}(B)(x)|_{x=0} = 0$ .

Set

$$F(A, B)(x) := (A(x), g(\mathcal{D}^1 A(x), \mathcal{D}^2[-\mathcal{P}\mathcal{A}A + \mathcal{P}B](x))).$$

Fix (small) open neighborhoods  $N_0, U_0 \subseteq \mathcal{C}^{1+s_0}(B^n(D);\mathbb{R}^{m_1}) \times \mathcal{C}^{s_0}(B^n(D);\mathbb{R}^{m_2})$  of  $(0,0)$ , to be chosen later. We take  $U_0 = U_0(N_0)$  small enough that  $F : U_0 \rightarrow N_0$ .

**LEMMA A.5.** *There exists an open neighborhood  $W_0 \subseteq N_0$  of  $(0,0)$  and a map  $G : W_0 \rightarrow U_0$  such that  $F(G(A, B)) = (A, B)$  and*

$$(A.16) \quad \begin{aligned} & \|G(A, B)\|_{\mathcal{C}^{1+s_0}(B^n(D);\mathbb{R}^{m_1}) \times \mathcal{C}^{s_0}(B^n(D);\mathbb{R}^{m_2})} \\ & \leq C \|(A, B)\|_{\mathcal{C}^{1+s_0}(B^n(D);\mathbb{R}^{m_1}) \times \mathcal{C}^{s_0}(B^n(D);\mathbb{R}^{m_2})}, \end{aligned}$$

where  $C$  does not depend on the choice of  $(A, B) \in W_0$ .

*Proof.* It is clear that  $F$  is a  $C^1$  map  $F : U_0 \subseteq \mathcal{C}^{1+s_0} \times \mathcal{C}^{s_0} \rightarrow N_0 \subseteq \mathcal{C}^{1+s_0} \times \mathcal{C}^{s_0}$  with  $F(0) = 0$  and  $dF(0) = I$ . The lemma now follows from the usual Inverse Function Theorem on Banach spaces.  $\square$

Let  $W_0$  be as in Lemma A.5 and set  $W := \{A : (A, 0) \in W_0\}$ . Note that  $W \subseteq \mathcal{C}^{1+s_0}(B^n(D);\mathbb{R}^{m_2})$  is an open neighborhood of 0. Taking  $G$  as in Lemma A.5 it is easy to see that  $G$  is of the form  $G(A, B) = (A, \tilde{G}(A, B))$ . We set

$$\mathcal{B}(A) := -\mathcal{P}\mathcal{A}A + \mathcal{P}\tilde{G}(A, 0).$$

It is clear that  $\mathcal{B}$  satisfies (A.12). By taking  $N_0$  small, we may take  $U_0$  and  $W$  as small as we like. Thus, because the range of  $G$  is contained in  $U_0$ , if  $N_0, U_0$ , and  $W$  are chosen to be sufficiently small we have  $\mathcal{B} : W \rightarrow N$ . Furthermore, by the choice of  $\mathcal{P}$  we have  $\mathcal{D}^1 \mathcal{B}(A)(0) = 0$ . Also, (A.13) follows from (A.16) and the continuity of  $\mathcal{P}$ .

It remains to prove (A.14) and (A.15). For this, we use that we have the flexibility to take  $U_0$  and  $N_0$  as small as we like (though they must be chosen independent of  $s$ ).

Let  $\gamma = \gamma(\mathcal{E}_2) > 0$  be as in Proposition A.1. By taking  $N_0$  and  $U_0$  sufficiently small, and using the fact that  $g_1$  is smooth, we have for  $A \in W$ , every coefficient of the differential operator  $\mathcal{L} := (g_1(A(x), \mathcal{D}^1 \mathcal{B}(A)(x)) - g_1(0, 0)) \mathcal{D}_2$  has  $L^\infty$  norm  $\leq \gamma$ ; indeed, since  $W_0 \subseteq N_0$ , taking  $N_0$  small forces  $W_0$ , and therefore  $W$ , to be a small neighborhood of 0. Setting  $B = \mathcal{B}(A)$ , we will apply Proposition A.1 (with  $u = B$ ) to the equation

$$(A.17) \quad (\mathcal{E}_2 + \mathcal{L})B = g_1(A(x), \mathcal{D}^1 B(x)) \mathcal{D}_2 B(x) = -g_2(\mathcal{D}^1 A(x), \mathcal{D}^1 B(x)).$$

Let  $0 < \eta_1 < \eta_2 \leq D$ . We will show for  $s > s_2 \geq s_0$ ,  $A \in W$ ,  $B = \mathcal{B}(A)$ ,

$$(A.18) \quad \begin{aligned} A &\in \mathcal{C}^{1+s}(B^n(D); \mathbb{R}^{m_1}), \quad B \in \mathcal{C}^{2+s_2}(B^n(\eta_2); \mathbb{R}^{m_2}) \\ &\implies B \in \mathcal{C}^{2+s_2+\min\{\frac{1}{2}, s-s_2\}}(B^n(\eta_1); \mathbb{R}^{m_2}), \end{aligned}$$

with

$$(A.19) \quad \|B\|_{\mathcal{C}^{2+s_2+\min\{\frac{1}{2}, s-s_2\}}(B^n(\eta_1); \mathbb{R}^{m_2})} \leq C_{s, s_2, \eta_1, \eta_2},$$

where  $C_{s, s_2, \eta_1, \eta_2}$  can be chosen to depend on  $\|A\|_{\mathcal{C}^{1+s}(B^n(D); \mathbb{R}^{m_1})}$  and  $\|B\|_{\mathcal{C}^{2+s_2}(B^n(\eta_2); \mathbb{R}^{m_2})}$ , but not depend on  $A$  or  $B$  in any other way. It can depend on any other ingredient in the problem. (A.14) and (A.15) follow from (A.18) and (A.19) via a simple iteration. Thus we prove (A.18) and (A.19) which will complete the proof.

Since  $g_1$  and  $g_2$  are smooth, if  $A \in \mathcal{C}^{1+s}$  and  $B \in \mathcal{C}^{2+s_2}$ , we have  $g_1(A, \mathcal{D}^1 B) \in \mathcal{C}^{s_2+1}$  and  $g_2(\mathcal{D}^1 A, \mathcal{D}^1 B) \in \mathcal{C}^{\min\{s, s_2+1\}} \subseteq \mathcal{C}^{\min\{s, s_2+\frac{1}{2}\}}$  (see Lemma 5.8). Furthermore, we have

$$(A.20) \quad \|g_1(A, \mathcal{D}^1 B)\|_{\mathcal{C}^{s_2+1}}, \|g_2(\mathcal{D}^1 A, \mathcal{D}^1 B)\|_{\mathcal{C}^{\min\{s, s_2+\frac{1}{2}\}}} \leq C_{s, s_2, \eta_1, \eta_2},$$

where  $C_{s, s_2, \eta_1, \eta_2}$  is as above; in particular, the estimate on  $g_1(A, \mathcal{D}^1 B)$  in (A.20) shows that the coefficients of  $\mathcal{L}$  are in  $\mathcal{C}^{s_2+1}$  with  $\mathcal{C}^{s_2+1}$  norms bounded by  $C_{s, s_2, \eta_1, \eta_2}$ , where  $C_{s, s_2, \eta_1, \eta_2}$  is as above. Applying Proposition A.1 to (A.17) with  $M = 2$ ,  $s_0 = s_2$ ,  $s = \min\{s, s_2 + \frac{1}{2}\}$ , and  $\epsilon_0 = \frac{1}{2}$ , and using the estimate on  $g_2(\mathcal{D}^1 A, \mathcal{D}^1 B)$  in (A.20), (A.18) and (A.19) follow, completing the proof.

**A.4. An elliptic operator.** In this section, we discuss a particular first order, overdetermined, constant coefficient, linear, elliptic operator which is needed in this paper. For a function  $A = (A_1, \dots, A_n) \in C^\infty(\mathbb{R}^n; \mathbb{R}^n)$  we define

$$\mathcal{E}A := \left( \left( \frac{\partial}{\partial t_j} A_k - \frac{\partial}{\partial t_k} A_j \right)_{1 \leq j < k \leq n}, \sum_{j=1}^n \frac{\partial}{\partial t_j} A_j \right).$$

LEMMA A.6.  $\mathcal{E}$  is elliptic.

*Proof.* It is easy to compute  $\mathcal{E}^* \mathcal{E}$  directly to see

$$\mathcal{E}^* \mathcal{E} A = - \sum_{j=1}^n \frac{\partial^2}{\partial t_j^2} A,$$

and the result follows.  $\square$

A more abstract way to see Lemma A.6 is as follows. We identify  $A$  with the 1-form  $A = A_1 dt_1 + A_2 dt_2 + \dots + A_n dt_n$ . Then,

$$dA = \sum_{1 \leq j < k \leq n} \left( \frac{\partial}{\partial t_j} A_k - \frac{\partial}{\partial t_k} A_j \right) dt_j \wedge dt_k, \quad \delta A = - \sum_{j=1}^n \frac{\partial}{\partial t_j} A_j,$$

where  $\delta$  denotes the codifferential on  $\mathbb{R}^n$ . Hence,  $\mathcal{E}$  can be written as  $\mathcal{E}A = (dA, -\delta A)$ , and therefore  $\mathcal{E}^* \mathcal{E} = d\delta + \delta d$ . I.e.,  $\mathcal{E}^* \mathcal{E}$  is the Laplace-de Rham operator acting on 1-forms, and is therefore elliptic.

UNIVERSITY OF WISCONSIN-MADISON, DEPARTMENT OF MATHEMATICS, 480 LINCOLN DR., MADISON, WI 53706

*E-mail:* street@math.wisc.edu

---

REFERENCES

---

- [DK81] D. M. DeTurck and J. L. Kazdan, Some regularity theorems in Riemannian geometry, *Ann. Sci. École Norm. Sup. (4)* **14** (1981), no. 3, 249–260.
- [FSC86] C. L. Fefferman and A. Sánchez-Calle, Fundamental solutions for second order subelliptic operators, *Ann. of Math. (2)* **124** (1986), no. 2, 247–272.
- [Gui08] V. Guillemin, Lecture notes, <http://math.mit.edu/~arita/18.101/>, prepared by Ana Rita Pires, accessed: 2017-05-02.
- [Mal69] B. Malgrange, Sur l'intégrabilité des structures presque-complexes, *Symposia Mathematica, Vol. II (INDAM, Rome, 1968)*, Academic Press, London, 1969, pp. 289–296.
- [NSW85] A. Nagel, E. M. Stein, and S. Wainger, Balls and metrics defined by vector fields. I. Basic properties, *Acta Math.* **155** (1985), no. 1-2, 103–147.

- [Nic07] L. I. Nicolaescu, *Lectures on the Geometry of Manifolds*, 2nd ed., World Scientific Publishing, Hackensack, NJ, 2007.
- [SS18] B. Stovall and B. Street, Coordinates adapted to vector fields: canonical coordinates, *Geom. Funct. Anal.* **28** (2018), no. 6, 1780–1862.
- [Str11] B. Street, Multi-parameter Carnot-Carathéodory balls and the theorem of Frobenius, *Rev. Mat. Iberoam.* **27** (2011), no. 2, 645–732.
- [Str20] \_\_\_\_\_, Sub-Hermitian geometry and the quantitative Newlander-Nirenberg theorem, *Adv. Math.* **368** (2020), 107137, 87.
- [Str21] \_\_\_\_\_, Coordinates adapted to vector fields III: real analyticity, *Asian J. Math.* (to appear).
- [TW03] T. Tao and J. Wright,  $L^p$  improving bounds for averages along curves, *J. Amer. Math. Soc.* **16** (2003), no. 3, 605–638.
- [Tay11] M. E. Taylor, *Partial Differential Equations III. Nonlinear Equations*, 2nd ed., *Appl. Math. Sci.*, vol. 117, Springer-Verlag, New York, 2011.
- [Tri06] H. Triebel, *Theory of Function Spaces. III, Monogr. Math.*, vol. 100, Birkhäuser Verlag, Basel, 2006.