Balls and metrics defined by vector fields I: Basic properties

by

ALEXANDER NAGEL[1], ELIAS M. STEIN[2] and STEPHEN WAINGER[1]

University of Wisconsin Madison, WI, U.S.A.

Princeton University Princeton, NJ, U.S.A.

University of Wisconsin Madison, WI, U.S.A.

Contents

Introduction

In this paper we study the basic properties of certain balls and metrics that can be naturally defined in terms of a given family of vector fields. As an application, we then use these properties to obtain estimates for the kernels of approximate inverses of some non-elliptic partial differential operators, such as Hörmander's sum of squares. Some

(1) Supported in part by an NSF grant at the University of Wisconsin, Madison.
(2) Supported in part by an NSF grant at Princeton University.

of the properties that we establish were announced earlier in [NSW]. In that paper, we also announced applications of these balls and metrics to problems concerning the boundary behavior of holomorphic functions in domains of finite type. We shall give details of these applications in a second paper.

To motivate our discussion, consider the following basic example. Let $\Omega\subset\mathbf{R}^N$ be a connected open set, and let $X_0, X_1, \dots, X_p$ be C^∞ real vector fields defined in a neighborhood of $\bar{\Omega}$. Suppose that there is an integer m so that these vector fields $X_0, X_1, \dots, X_p$, together with their commutators of length at most m span $\mathbf{R}^N$ at each point of $\bar{\Omega}$. In this case we say that $X_0, \dots, X_p$ are of *type* m on Ω. Let

$$X^{(1)} = \{X_0, \dots, X_p\}$$
$$X^{(2)} = \{[X_0, X_1], \dots, [X_{p-1}, X_p]\}, \quad \text{etc.}$$

so that the components of $X^{(k)}$ are the commutators of length k. Let $Y_1, \dots, Y_q$ be some enumeration of the components of $X^{(1)}, \dots, X^{(m)}$. If Y_i is an element of $X^{(j)}$, we say Y_i has formal degree $d(Y_i)=j$.

Now define a metric ϱ on Ω by requiring that $\varrho(x_0, x_1)<\delta$ if and only if there is an absolutely continuous map $\varphi\colon[0,1]\to\Omega$ with $\varphi(0)=x_0$, $\varphi(1)=x_1$, and for almost all $t\in[0,1]$

$$\varphi'(t)=\sum_{j=1}^{q} a_j(t)\, Y_j(\varphi(t))$$

with $|a_j(t)|<\delta^{d(Y_j)}$. There is then a corresponding family of balls on Ω given by

$$B(x,\delta)=\{y\in\Omega\,|\,\varrho(x,y)<\delta\}.$$

These balls reflect the non-isotropic nature of the vector fields $X_0, \dots, X_p$ and their commutators. A ball $B(x_0, \delta)$ is essentially of size δ in the directions from x_0 specified by $X_0, \dots, X_p$, but only of size δ^2 in the directions given by commutators of length 2, of size δ^3 in the directions given by commutators of length 3, etc. Such balls play an important role in the study of boundary behavior of holomorphic functions (see [St2]), and, as we shall see, in studying the hypoelliptic operator $X_0^2+X_1^2+\dots+X_p^2$.

Our object in this paper is to study the basic properties of these balls. For example we obtain alternate descriptions of these balls in terms of exponential maps, and estimates on their volume. We also study the behavior of the balls under suitable mappings of the underlying space, and this allows us to estimate kernels for parametrices for $X_0^2+X_1^2+\dots+X_p^2$.

We would like to give some explanation for much of the formalism that follows. Suppose we have N vector fields, $Y_1, \dots, Y_N$, on $\mathbf{R}^N$ such that at every point of $\mathbf{R}^N$, $Y_1, \dots, Y_N$ formed a basis for the tangent space to $\mathbf{R}^N$. Then it would be natural to perform calculations near a given point x in terms of canonical coordinates at x. This means that if y is near x and

$$y = \exp(a_1 Y_1 + \dots + a_N Y_N)(x)$$

we assign to y the coordinates $(a_1, \dots, a_N)$. (See the appendix for a more complete discussion.)

In our situation, we have q vector fields $Y_1, Y_2, \dots, Y_q$ where in general $q > N$ and at each point of $\mathbf{R}^N$ many different subsets of $\{Y_1, \dots, Y_q\}$ form a basis for the tangent space to $\mathbf{R}^N$. One of our main difficulties was to find a good choice for a subcollection $Y_{i_1}, \dots, Y_{i_N}$ of the $Y_1, \dots, Y_q$ to choose as a basis for performing calculations. Roughly speaking, for a given x and $\varrho(x, y) \approx \delta$, we choose $Y_{i_1}, \dots, Y_{i_N}$ such that

$$\text{(i)} \qquad \delta^{\deg Y_{i_1} + \dots + \deg Y_{i_N}} \det(Y_{i_1}, \dots, Y_{i_N})(x)$$

is maximal. The motivation for this choice of $Y_{i_1}, \dots, Y_{i_N}$ arises from the fact that the volume of the image of the ball

$$\{a \mid |a_{i_j}| < \delta^{\deg Y_{i_j}},\ j = 1, 2, \dots, N)\}$$

under the exponential map corresponding to $Y_{i_1}, \dots, Y_{i_N}$ is given by (i).

One feature of our work is the algebraic nature of our estimates. Suppose we have an N-tuple $I = (i_1, \dots, i_N)$ and $\det(Y_{i_1}, \dots, Y_{i_N}) \neq 0$ near x. Then for any j, we may write (near x)

$$Y_j = \sum_{l=1}^{N} a_{j,I}^l Y_{i_l}.$$

We shall let $A_s^p = A_{s,I}^p$ denote the submodule of C^∞ functions generated by products

$$a_{j_1}^{l_1} \dots a_{j_k}^{l_k}$$

where $k \leq s$ and

$$p \leq \deg Y_{l_1} + \dots + \deg Y_{l_k} - \deg Y_{j_1} + \dots + \deg Y_{j_k}$$

where $l_1, \dots, l_k$ are arbitrary and $j_1, \dots, j_k$ are in the N-tuple I. We may obtain estimates on functions by showing that they belong to appropriate modules A_s^p. This will require considerable computation. We will need to show first of all that if I is chosen so that the quantity (i) is maximal, then

$$\text{(ii)} \qquad |a^l_{j_k}(x)| \leq C\delta^{\deg Y_{i_l} - \deg Y_{j_k}}.$$

We shall want (ii) to hold for y's near x. To do this we shall need to estimate derivatives of $a^l_{j_k}$ at x. Thus we will want to show derivatives of elements of A^p_s are in appropriate modules. It will be particularly important for us to estimate also $\det(Y_{j_1}, \ldots, Y_{j_N})$ for arbitrary N-tuples $(j_1, \ldots, j_N)$. Thus we shall also have to see to which modules these functions and their derivatives belong. It will also be important for us to estimate from below the set on which various exponential maps are one-to-one. (This is important, for example, in computing volumes of balls.) It turns out that this is quite intricate, but that at least we can get information on the set where the exponential map is locally one-to-one by studying $\det(Y_{j_1}, \ldots, Y_{j_N})$.

The study of geometric properties of vector fields and their commutators has a very long history. Carathéodory [C] was the first to prove that if commutators of sufficiently high order m of a family of vector fields span at every point, then any two points can be joined by a piecewise smooth curve whose tangents belong to the family of vector fields. (See also the paper of W. L. Chow [Ch].) In proving hypoellipticity of certain operators, Hörmander [H] studied differentiability along non-commuting vector fields, and used the techniques of exponential mappings, and the Campbell-Hausdorff formula. The case of vector fields of type 2 were studied in [NS], sections 5 and 14. Balls reflecting commutation properties of vector fields have also been studied by Folland and Hung [FH], by Fefferman and Phong [FP], and more recently by Sanchez [Sa], who has independently obtained some of our results.

We would like to thank Joel Robbin for helpful conversations about exponential maps and the Campbell-Hausdorff formula.

Chapter 1. Definitions and statement of results

§ 1. The basic metric ϱ

Throughout this paper, we shall be concerned with the following situation. Let $\Omega \subset \mathbf{R}^N$ be a connected open set, and suppose $Y_1, \ldots, Y_q$ are C^∞ real vector fields defined on a neighborhood of $\bar{\Omega}$. We suppose that each vector field Y_j has associated a formal degree $d_j \geq 1$, where d_j is an integer. We now make the following hypotheses:

(a) For each j and k we can write

$$[Y_j, Y_k] = \sum_{d_l \leq d_j + d_k} c_{jk}^l(x)\, Y_l \tag{1}$$

where $c_{jk}^l \in C^\infty(\bar{\Omega})$.

(b) For each $x \in \bar{\Omega}$, the vectors $Y_1(x), \ldots, Y_q(x)$ span $\mathbf{R}^N$.

A basic example is the one discussed in the introduction of a collection of vector fields $X_0, X_1, \ldots, X_p$ of finite type m. In that case, property (a) follows from the Jacobi identity, and property (b) is immediate. This example will be used to study the differential operator $X_0{}^2 + X_1{}^2 + \ldots + X_p{}^2$, and its variants $X_0{}^2 + X_1{}^2 + \ldots + X_p{}^2 + \frac{1}{2}\Sigma_{j,k}\, c_{jk}[X_j, X_k]$ considered in [RS].

In order to study the modified operator

$$X_1{}^2 + X_2{}^2 + \ldots + X_p{}^2 + X_0$$

we would again let $Y_1, \ldots, Y_q$ be some enumeration of the components of $X^{(1)}, \ldots, X^{(m)}$. However this time the vectors $X_1, \ldots, X_q$ would have degree 1, while X_0 would have degree 2, and higher length commutators would be weighted accordingly. We shall return to these examples later.

Returning to the general situation of vector fields $Y_1, \ldots, Y_q$ on Ω satisfying (a) and (b), we can define a metric on Ω in the following way:

Definition 1.1. Let $C(\delta)$ denote the class of absolutely continuous mappings $\varphi\colon [0,1] \to \Omega$ which almost everywhere satisfy the differential equation

$$\varphi'(t) = \sum_{j=1}^{q} a_j(t)\, Y_j(\varphi(t)) \tag{2}$$

with

$$|a_j(t)| < \delta^{d_j}. \tag{3}$$

Then define

$$\varrho(x, y) = \inf\{\delta > 0 \mid \exists \varphi \in C(\delta) \text{ with } \varphi(0) = x,\ \varphi(1) = y\}. \tag{4}$$

We then have the following simple

PROPOSITION 1.1. *ϱ is a metric on Ω. If $m = \max d_j$, and if $K \subset\subset \Omega$ is any compact set, there are constants C_1, C_2 so that if $x, y \in K$*

$$C_1|x-y| \leq \varrho(x, y) \leq C_2|x-y|^{1/m}.$$

Proof. It is clear from the definition that ϱ is a metric. Let $K \subset\subset \Omega$ be an arcwise connected compact set. There is a constant C so that if $x, y \in K$, there is an absolutely continuous function $\varphi: [0,1] \to \Omega$ with $\varphi(0)=x$, $\varphi(1)=y$ and $|\varphi'(t)| \le C|x-y|$ for all t. Since the vector fields $Y_1, \ldots, Y_q$ span at each point, we can write

$$\varphi'(t) = \sum_{j=1}^{q} b_j(t)\, Y_j(\varphi(t))$$

with $|b_j(t)| \le C'|\varphi'(t)| \le C''|x-y| = C''(|x-y|^{1/d_j})^{d_j}$. Since $d_j \le m$ it follows that

$$\varrho(x,y) \le C|x-y|^{1/m}.$$

Conversely, if $x, y \in K$ and $\varrho(x,y)=\delta$, there exists $\varphi \in C(2\delta)$ with $\varphi(0)=x$ and $\varphi(1)=y$, and $\varphi'(t) = \sum_{j=1}^{q} a_j(t)\, Y_j(\varphi(t))$ with $|a_j(t)| \le (2\delta)^{d_j}$. Since the components of every Y_j are uniformly bounded on $\bar{\Omega}$, it follows that

$$|\varphi'(t)| \le C \sum_{j=1}^{q} (2\delta)^{d(Y_j)} \le C'\delta.$$

Hence

$$|x-y| = \left| \int_0^1 \varphi'(t)\, dt \right| \le C'\delta. \qquad \text{Q.E.D.}$$

It follows from the proposition that the metric $\varrho: \Omega \times \Omega \to [0, \infty)$ is continuous. We can define a family of balls $B(x,\delta)$ on Ω by

$$B(x,\delta) = \{y \in \Omega \mid \varrho(x,y) < \delta\}. \tag{5}$$

§ 2. General families of balls

In studying families of balls, we are primarily concerned with those properties that are essential for the covering lemma used in the proof of the Hardy-Littlewood maximal theorem (see [St1], Chapter 1). We begin by listing these properties for a family of balls defined on an open subset $\Omega \subset \mathbf{R}^N$.

For each $x \in \Omega$ and each δ with $0 < \delta \le \delta_0$ suppose we are given a set $B(x,\delta) \subset \Omega$, called a ball with center x and radius δ. We are then concerned with the following properties:

(i) $B(x,\delta)$ is open, and if $0<\delta\le\delta_0$, $B(x,\delta) = \bigcup_{s<\delta} B(x,s)$.

(ii) $\bigcap_{s>0} B(x,s) = \{x\}$.

(iii) For every compact set $K \subset\subset \Omega$ there is a constant C so that if $x_1, x_2 \in K$, if $\delta_1 \leq \delta_2 \leq (1/C)\delta_0$, and if

$$B(x_1, \delta_1) \cap B(x_2, \delta_2) \neq \emptyset$$

then $B(x_1, \delta) \subset B(x_2, C\delta_2)$.

(iv) For every compact $K \subset\subset \Omega$ there is a constant C so that if $x \in K$ and if $\delta < \frac{1}{2}\delta_0$

$$|B(x, 2\delta)| \leq C|B(x, \delta)|.$$

Here, and in the rest of the paper, $|E|$ denotes the Lebesgue measure of a measurable set $E \subset \mathbf{R}^N$.

Now to any family of balls satisfying (i), (ii), and (iii), we can associate a quasi-distance $\varrho\colon \Omega \times \Omega \to [0, \infty]$ defined by

$$\varrho(x, y) = \begin{cases} \inf\{\delta > 0 \mid y \in B(x, \delta)\} & \text{if } y \in B(x, \delta_0) \\ +\infty & \text{if } y \notin B(x, \delta_0) \end{cases} \tag{6}$$

and the family of balls $\{B(x, \delta)\}$ can be recovered from this quasi-distance, since

$$B(x, \delta) = \{y \in \Omega \mid \varrho(x, y) < \delta\}. \tag{5}$$

The function ϱ satisfies the following conditions:

(i′) For every $x \in \Omega$ the set $\{y \in \Omega \mid \varrho(x, y) < \delta\}$ is open.

(ii′) $\varrho(x, y) = 0$ if and only if $x = y$.

(iii′) For every compact set $K \subset\subset \Omega$ there is a constant C so that if $x, y, z \in K$,

$$\varrho(x, y) \leq C[\varrho(x, z) + \varrho(y, z)].$$

(Note that by choosing $z = x$, (iii′) implies $\varrho(x, y) \leq C\varrho(y, x)$.)

Conversely, if $\varrho\colon \Omega \times \Omega \to [0, \infty]$ is a function satisfying (i′), (ii′), and (iii′) then the balls defined by equation (5) satisfy conditions (i), (ii), and (iii). In particular this is true if ϱ is actually a metric.

We next introduce a notion of equivalence of families of balls. We say that two functions $\varrho_1, \varrho_2\colon \Omega \times \Omega \to [0, \infty]$ are *equivalent* if for every compact set $K \subset\subset \Omega$ there is a constant C so that if $X_1, X_2 \in K$

$$\begin{aligned} &\varrho_1(x_1, x_2) \leq C\varrho_2(x_1, x_2), \quad \text{and} \\ &\varrho_2(x_1, x_2) \leq C\varrho_1(x_1, x_2). \end{aligned} \tag{7}$$

We say that ϱ_1 and ϱ_2 are *locally equivalent* if for each $x_0 \in \Omega$ there is an open neighborhood U containing x_0 so that ϱ_1 and ϱ_2 are equivalent on U.

If ϱ_1 and ϱ_2 are quasi-distances, and B_1, B_2 are the corresponding families of balls, then equivalence means exactly that for every compact set $K \subset\subset \Omega$ there is a constant C so that for $x \in K$

$$\begin{aligned} &B_1(x,\delta) \subset B_2(x, C\delta), \quad \text{and} \\ &B_2(x,\delta) \subset B_1(x, C\delta). \end{aligned} \tag{8}$$

It is clear that if a family of balls satisfies conditions (i), (ii), (iii), and (iv), so will any equivalent family of balls.

§ 3. Volumes of balls

The function ϱ defined in § 1 is a metric, and so the corresponding family of balls $\{B(x,\delta)\}$ satisfies conditions (i), (ii), and (iii) of § 2. One of the main goals of this paper is to show that the balls defined by the metric ϱ also satisfy condition (iv), the "doubling property" for volume. In fact, we can give very explicit estimates for the volumes of the balls $B(x,\delta)$ in the following way:

For each N-tuple of integers $I=(i_1, \ldots, i_N)$ with $1 \leq i_j \leq q$, set

$$\lambda_I(x) = \det(Y_{i_1}, \ldots, Y_{i_N})(x). \tag{9}$$

(If $Y_{i_j} = \Sigma_{k=1}^N a_{jk}(x)(\partial/\partial x_k)$, then $\det(Y_{i_1} \ldots Y_{i_N})(x) = \det(a_{jk}(x))$. We also set

$$d(I) = d_{i_1} + \ldots + d_{i_N} \tag{10}$$

and then we define

$$\Lambda(x,\delta) = \sum_I |\lambda_I(x)|\, \delta^{d(I)} \tag{11}$$

where the sum is over all N-tuples. With this notation we can now state our main result on the volumes of the balls $B(x,\delta)$.

THEOREM 1. *For every compact set $K \subset\subset \Omega$ there are constants C_1 and C_2 so that for all $x \in K$*

$$0 < C_1 \leq \frac{|B(x,\delta)|}{\Lambda(x,\delta)} \leq C_2 < +\infty.$$

Since Λ is a polynomial in δ of fixed degree, it follows immediately from Theorem 1 that there is a constant C so that for $x \in K$

$$|B(x, 2\delta)| \leq C|B(x, \delta)|$$

and thus condition (iv) is satisfied.

§ 4. Equivalent pseudo-distances

It seems difficult to compute or estimate the distance $\varrho(x, y)$ directly from the definition. Thus a major ingredient in the proof of Theorem 1, and another major object of the paper, is to understand the relationship between the metric ϱ and various other metrics or quasi distances which may be easier to estimate.

Our first alternative pseudo-distance is similar to ϱ except that we allow only *constant* linear combinations of the vectors $Y_1, \ldots, Y_q$. Thus for $\delta > 0$ let $C_2(\delta)$ denote the class of smooth curves $\varphi: [0, 1] \to \Omega$ such that

$$\varphi'(t) = \sum_{j=1}^{q} a_j Y_j(\varphi(t)) \tag{12}$$

with $|a_j| < \delta^{d_j}$.

Define

$$\varrho_2(x, y) = \inf\{\delta > 0 \mid \exists \varphi \in C_2(\delta) \text{ with } \varphi(0) = x,\ \varphi(1) = y\}. \tag{13}$$

It is not necessarily true that $\varrho_2(x, y)$ is finite for every $(x, y) \in \Omega \times \Omega$. (For example, $\varrho_2(x, y)$ can be infinite if Ω is a non-convex subset of $\mathbf{R}^2$, and $Y_1 = \partial/\partial x_1$, $Y_2 = \partial/\partial x_2$.) We shall prove however:

THEOREM 2. *ϱ_2 is locally equivalent to ϱ.*

While the curves of class $C_2(\delta)$ are somewhat easier to study in that only constant linear combinations of $Y_1, \ldots, Y_q$ are allowed, there is still the disadvantage that the number q is in general much larger than the dimension N. We would like to single out N of the vector fields $Y_1, \ldots, Y_q$, say $Y_{i_1}, \ldots, Y_{i_N}$, so that the exponential mapping

$$(u_1, \ldots, u_N) \in \mathbf{R}^N \mapsto \exp\left(\sum_{j=1}^{N} u_j Y_{i_j}\right)(x_0) \tag{14}$$

is appropriate for performing calculations. (The definition and basic properties of such exponential mappings are recalled in the appendix, Chapter 4.) Roughly, we would like the ball $B(x, \delta)$ to be essentially the image under the mapping (14) of the set

$$\{(u_1, \dots, u_N) \in \mathbf{R}^N \mid |u_j| < \delta^{d_{i_j}}\}.$$

It turns out that one cannot make a single choice of vector fields which works for all x and δ, or even for a fixed x and all δ. Rather one must choose the N-tuple depending on both x and δ.

Thus for each N-tuple $I=(i_1, \dots, i_N)$, let $C_3(\delta, I)$ denote the class of smooth curves $\varphi: [0, 1] \to \Omega$ such that

$$\varphi'(t) = \sum_{j=1}^{N} a_j Y_{i_j}(\varphi(t)) \tag{15}$$

with $|a_j| < \delta^{d(Y_{i_j})}$. Set $C_3(\delta) = \bigcup_I C_3(\delta, I)$, and define

$$\varrho_3(x, y) = \inf\{\delta > 0 \mid \exists \varphi \in C_3(\delta) \text{ with } \varphi(0) = x,\ \varphi(1) = y\}. \tag{16}$$

In studying ϱ_3, we choose an N-tuple I depending on both x and δ. Roughly, the idea is to choose I so that

$$\delta^{d(I)} |\lambda_I(x)|$$

is as large as possible. Then (14) will be an appropriate exponential mapping. We will be able to prove

THEOREM 3. *ϱ_3 is locally equivalent to ϱ and ϱ_2.*

We will also show that with the appropriate choice of N-tuple, the mapping (14) is one to one on a sufficiently large set, and this will finally enable us to obtain the volume estimates of Theorem 1.

In the case that the family of vector fields $Y_1, \dots, Y_q$ is simply an enumeration of the commutators of $X_0, \dots, X_p$ with the degree of Y_i equaling the length of the commutator, there is another natural metric which we can define.

Thus let $C_4(\delta)$ denote the class of absolutely continuous mappings $\varphi: [0, 1] \to \Omega$ which almost everywhere satisfy

$$\varphi'(t) = \sum_{j=1}^{p} a_j(t) X_j(\varphi(t)) \quad \text{with } |a_j(t)| < \delta. \tag{17}$$

Note that curves in the class $C_4(\delta)$ point *only* in the directions of the original vector fields $X_0, \dots, X_p$. Define

$$\varrho_4(x,y)=\inf\{\delta>0 \mid \exists\varphi\in C_2(\delta) \text{ with } \varphi(0)=x,\ \varphi(1)=y\}. \tag{18}$$

$\varrho_4\colon \Omega\times\Omega\to[0,\infty)$ is again a metric, although it is not clear *a pri*ori that $\varrho_4(x,y)$ is finite for every $(x,y)\in\Omega\times\Omega$. The fact that ϱ_4 is finite follows because the commutators of the vector fields $X_0, \dots, X_p$ span $\mathbf{R}^N$ at each $x\in\Omega$. This was first proved by Carathéodory. We shall prove:

THEOREM 4. *ϱ_4 is equivalent to ϱ.*

§ 5. Estimates for certain kernels

The balls and metrics we have defined can be used to estimate the sizes of certain kernels and their derivatives. These kernels were constructed by Rothschild and Stein in [RS]. Let us briefly describe the setting. Given vector fields $X_0, \dots, X_p$ on Ω of type m, then according to [RS] we can add additional variables $(t_1, \dots, t_s)\in\mathbf{R}^s$ and form new vector fields

$$\tilde{X}_j=X_j+\sum_{l=1}^{s} a_{jl}(x,t)\frac{\partial}{\partial t_l}. \tag{19}$$

These vector fields will again be of type m on an appropriate set in $\mathbf{R}^N\times\mathbf{R}^s$, but in addition, they are "free up to step m"; i.e. their commutators of length at most m satisfy no linear relations except those dictated by antisymmetry and the Jacobi identity.

The vector fields $\tilde{X}_1, \dots, \tilde{X}_p$ give rise to a metric $\tilde{\varrho}$ and a family of balls $\tilde{B}((x,t),\delta)$. If $k(x,t;y,s)$ is a kernel in (x,t) space, we can form a kernel $Rk(x,y)$ on $\Omega\times\Omega$ by setting

$$Rk(x,y)=\int k(x,0;y,s)\,\varphi(s)\,ds \tag{20}$$

where $\varphi(s)$ is C^∞, supported in $|s|\leqslant 1$, and identically 1 near the origin. We want to be able to make estimates on the size of Rk and its derivatives in terms of corresponding estimates on k. We have, for example

THEOREM 5. *Suppose*

$$|k(x,t;y,s)|\leqslant C\frac{|\tilde{\varrho}(x,t;y,s)|^{\alpha}}{|\tilde{B}[(x,t);\tilde{\varrho}(x,t;y,s)]|}.$$

Then

$$|Rk(x,y)| \leq C' \int_{\varrho(x,y)}^{1} \frac{r^{\alpha-1}}{|B(x,r)|}\,dr.$$

Moreover, if in addition

$$|\tilde{X}_{i_1}\tilde{X}_{i_2}\dots\tilde{X}_{i_j}k| \leq C\frac{|\tilde{\varrho}(x,t;y,s)|^{\alpha-j}}{|\tilde{B}[(x,t;\tilde{\varrho}(x,t,y,s)]|},$$

then

$$|X_{i_1}\dots X_{i_j}Rk| \leq C\int_{\varrho(x,y)}^{1} \frac{r^{\alpha-j-1}}{B(x,r)}\,dr.$$

The above estimates hold uniformly on compact subsets of $\Omega\times\Omega$. An interesting case arises when we consider

$$\Delta = X_0^{\,2}+\dots+X_p^{\,2}$$

or

$$\mathcal{H} = X_0+X_1^{\,2}+\dots+X_p^{\,2}$$

on Ω. Let $D(x,y)$ and $H(x,y)$ denote the fundamental solutions of Δ and $\mathcal{H}$ respectively constructed in [RS]. We then have

COROLLARY. *If* $N\geqq 3$

$$|D(x,y)| \leq C\frac{\varrho^2(x,y)}{|B(x,\varrho(x,y)|}.$$

If $N\geqq 2$

$$|X_{i_1}\dots X_{i_j}D| \leq C\frac{\varrho^{2-j}(x,y)}{|B(x,\varrho(x,y))|}.$$

If $N\geqq 2$

$$|H(x,y)| \leq C\frac{\varrho^2(x,y)}{|B(x,\varrho(x,y))|},$$

and if $i_s\neq 0,\ 1\leqq s\leqq j,$

$$X_{i_1}\dots X_{i_j}X_0^l H \leq C\frac{\varrho^{2-j-2l}}{|B(x,\varrho(x,y))|}.$$

The above estimates hold uniformly for x, y in compact subsets of $\Omega\times\Omega$. The corollary follows from the estimates of [RS], where the estimates for D and H are derived in the case of the free vector fields.

Remark. The assumption $N\geqslant 3$ in the corollary is necessary as we can see by considering the ordinary Laplacian in $\mathbf{R}^2$.

Chapter II. Structure of the balls

§ 1. Algebraic preliminaries

Our object in this section is to develop some algebraic machinery associated to a family of vector fields. Thus suppose $\Omega\subset\mathbf{R}^N$ is a connected open set, and suppose $\{Y_j\}$ is a (possibly infinite) collection of real C^∞ vector fields on Ω. We suppose that each vector field Y_j in this list has associated a formal degree $d_j\geqslant 1$, so that d_j is an integer. We also suppose that the number of vector fields with $d_j\leqslant M$ is finite for every M. Our fundamental hypothesis is that for all j and k we can write

$$[Y_j, Y_k] = \sum_{d_l\leqslant d_j+d_k} c_{jk}^l(x)\, Y_l \tag{1}$$

where the sum, of course, is finite, and $c_{jk}^l\in C^\infty(\bar{\Omega})$.

For any s-tuple of positive integers $J=(j_1, \dots, j_s)$ we let

$$d(J) = d_{j_1}+\dots+d_{j_s} \tag{2}$$

be the "degree" of J. We also let

$$Y_J = Y_{j_1} Y_{j_2} \dots Y_{j_s} \tag{3}$$

be the corresponding sth-order differential operator, and we let

$$Y_{[J]} = [Y_{j_1}, [Y_{j_2}, [\dots, [Y_{j_{s-1}}, Y_{j_s}]\dots] \tag{4}$$

be the corresponding commutator of length s. If $I=(i_1, \dots, i_N)$ is an N-tuple of positive integers, we let

$$\lambda_I(x) = \det(Y_{i_1}, \dots, Y_{i_N})(x) \tag{5}$$

and

$$\Omega_I = \{x\in\Omega \mid \lambda_I(x)\neq 0\}. \tag{6}$$

Ω_I is then a possibly empty open subset of Ω, and on Ω_I, the vectors $Y_{i_1}, \ldots, Y_{i_N}$ are a basis for the tangent space at every point.

LEMMA 2.1. *For every s-tuple J of positive integers, we can write*

$$Y_{[J]} = \sum_{d_I \leq d(J)} c^I_{[J]}(x)\, Y_I \tag{7}$$

where the sum is finite, and $c^I_{[J]} \in C^\infty(\bar{\Omega})$.

Proof. This follows from (1) by induction, and the formula

$$[X_1, fX_2] = (X_1 f)\, X_2 + f[X_1, X_2],$$

if X_1, X_2 are vector fields, and $f \in C^\infty(\Omega)$.

Now let $I = (i_1, \ldots, i_N)$ be an N-tuple of positive integers, and suppose $\Omega_I \neq \emptyset$. Then since $Y_{i_1}, \ldots, Y_{i_N}$ form a basis, on Ω_I we can write for any j

$$Y_j = \sum_{l=1}^{N} a^l_{j,I}(x)\, Y_{i_l} \tag{8}$$

where $a^l_{j,I} \in C^\infty(\Omega_I)$. Usually the N-tuple I will be understood, and we will write

$$Y_j = \sum_{l=1}^{N} a^l_j(x)\, Y_{i_l}. \tag{8$'$}$$

More generally, for any s-tuple $J = (j_1, \ldots, j_s)$ of positive integers, we can write

$$Y_{[J]} = \sum_{l=1}^{N} a^l_{J,I}(x)\, Y_{i_l} = \sum_{l=1}^{N} a^l_J(x)\, Y_{i_l} \tag{9}$$

where $a^l_J \in C^\infty(\Omega_I)$.

The coefficients a^l_j can be expressed in terms of the determinants $\{\lambda_J\}$. In fact, if we solve equation (8) or (8′) for the a^l_j by using Cramer's rule we obtain

LEMMA 2.2. *On the set Ω_I we have*

$$a^l_j(x) = \lambda_J(x)/\lambda_I(x) \tag{10}$$

where the N tuple J is obtained from I by replacing Y_{i_l} by Y_j.

We shall have to deal with sums of products of the coefficients $\{a_j^l\}$, and we formalize this by introducing certain finitely generated modules of functions. For each N-tuple I, we can regard $C^\infty(\Omega_I)$ as a module over the ring $C^\infty(\Omega)$. For every integer p and every positive integer s, *define:* $A_{s,I}^p = A_s^p$ is the $C^\infty(\Omega)$ submodule of $C^\infty(\Omega_I)$ generated by all functions of the form

$$a_{j_1}^{l_1} \cdot a_{j_2}^{l_2} \cdot \ldots \cdot a_{j_k}^{l_k}$$

where $k \leqslant s$, and

$$p \leqslant (d_{i_{l_1}} + d_{i_{l_2}} + \ldots + d_{i_{l_k}}) - (d_{j_1} + d_{j_2} + \ldots + d_{j_k}).$$

There are clearly only finitely many such functions, so A_s^p is finitely generated. We also have inclusions

$$A_s^{p_1} \supset A_s^{p_2} \quad \text{if } p_1 \leqslant p_2 \tag{11}$$

$$A_{s_1}^p \subset A_{s_2}^p \quad \text{if } s_1 \leqslant s_2. \tag{12}$$

Also note that $a_j^l \in A_1^{d_{i_l} - d_j}$, and since $a_{i_l}^l = 1$, $1 \in A_1^0$.

We will obtain estimates for various quantities by showing that they belong to the modules $A_{s,I}^p$. To obtain some idea of this, note that if I is an N-tuple such that $\lambda_I(x_0)\,\delta^{d(I)} \geqslant \sup_J \lambda_J(x_0)\,\delta^{d(J)}$, then from (10) we see that

$$|a_j^l(x_0)| \leqslant \delta^{d_{i_l} - d_j}.$$

LEMMA 2.3. *If $f \in A_s^p$, then $Y_i f \in A_{s+1}^{p-d_i}$.*

Proof. By Leibnitz's rule, it suffices to show that

$$Y_i(a_k^l) \in A_2^{d_{i_l} - d_k - d_i}.$$

From equation (1) we have

$$\begin{aligned}[]
[Y_i, Y_k] &= \sum_{d_j \leqslant d_i + d_k} c_{ik}^j(x)\, Y_j \\
&= \sum_{l=1}^{N} \left(\sum_{d_j \leqslant d_i + d_k} c_{ik}^j(x)\, a_j^l(x) \right) Y_{i_l}.
\end{aligned} \tag{13}$$

On the other hand, we can also write

$$
\begin{aligned}
[Y_i, Y_k] &= \left[Y_i, \sum_{l=1}^{N} a_k^l(x)\, Y_{i_l} \right] \\
&= \sum_{l=1}^{N} \left(Y_i(a_k^l)\, Y_{i_l} + a_k^l(x)\, [Y_i, Y_{i_l}] \right) \\
&= \sum_{l=1}^{N} Y_i(a_k^l)\, Y_{i_l} + \sum_{l=1}^{N} a_k^l(x) \sum_{d_r \leq d_i + d_{i_l}} c_{i,i_l}^r(x)\, Y_r \\
&= \sum_{l=1}^{N} Y_i(a_k^l)\, Y_{i_l} + \sum_{l=1}^{N} \sum_{d_r \leq d_i + d_{i_l}} \sum_{s=1}^{N} a_k^l(x)\, c_{i,i_l}^r(x)\, a_r^s(x)\, Y_{i_s}.
\end{aligned} \tag{14}
$$

Since $Y_{i_1}, \ldots, Y_{i_N}$ are a basis on Ω_I we can equate coefficients of Y_{i_l} in (13) and (14), and obtain

$$
Y_i(a_k^l) = \sum_{d_j \leq d_i + d_k} c_{ik}^j(x)\, a_j^l(x) - \sum_{s=1}^{N} \sum_{d_r \leq d_i + d_{i_s}} a_k^s(x)\, c_{i,i_s}^r(x)\, a_r^l(x). \tag{15}
$$

In the sum on the right hand side of (15), we have

$$
d_{i_l} - d_j \geq d_{i_l} - d_i - d_k
$$

for the first sum, and

$$
\begin{aligned}
d_{i_s} + d_{i_l} - d_k - d_r &\geq d_{i_s} + d_{i_l} - d_k - d_i - d_{i_s} \\
&= d_{i_l} - d_i - d_k
\end{aligned}
$$

for the second sum. Thus the lemma is proved.

We can now use induction to prove:

LEMMA 2.4. *If $J=(j_1, \ldots, j_k)$ is a k-tuple of positive integers, then if $f \in A_s^p$, $Y_J f \in A_{s+k}^{p-d(J)}$.*

We also can obtain information about the coefficients a_J^l of the commutator $Y_{[J]}$ of equation (9).

LEMMA 2.5. $a_J^l \in A_1^{d_{i_l} - d(J)}$.

Proof. According to Lemma 2.1,

$$Y_{[J]}=\sum_{d_j\leq d(J)} c^j_{[J]}(x)\, Y_j$$

with $c^j_{[J]}\in C^\infty(\Omega)$. Hence on Ω_I we can write

$$Y_{[J]}=\sum_{l=1}^{N}\left(\sum_{d_j\leq d(J)} c^j_{[J]}(x)\, a^l_j(x)\right) Y_{i_l}$$

and so

$$a^l_J=\sum_{d_j\leq d(J)} c^j_{[J]}(x)\, a^l_j$$

and since $d_{i_l}-d_j\geq d_{i_l}-d(J)$, the lemma is proved.

We shall also need to estimate derivatives of the determinants $\{\lambda_J\}$, and we begin with the following general formula. Let $Z_j=\Sigma_{k=1}^N a_{jk}(x)\,(\partial/\partial x_k)$, $j=1,\dots,N$ be N C^1 vector fields, and let $T=\Sigma_{l=1}^N \beta_l(x)\,(\partial/\partial x_l)$ be another.

LEMMA 2.6.

$$T(\det(Z_1,\dots,Z_N))=(\nabla\cdot T)\det(Z_1,\dots,Z_N)+\sum_{j=1}^{N}\det(Z_1,\dots,Z_{j-1},[T,Z_j],Z_{j+1},\dots,Z_N)$$

where $\nabla\cdot T=\Sigma_{l=1}^N \partial\beta_l/\partial x_l$.

Proof.

$$\begin{aligned}[T,Z_j]&=\sum_{l=1}^{N}\sum_{k=1}^{N}\left[\beta_l(x)\frac{\partial}{\partial x_l},\, a_{jk}(x)\frac{\partial}{\partial x_k}\right]\\ &=\sum_{k=1}^{N}\left(\sum_{l=1}^{N}\beta_l(x)\frac{\partial a_{j,k}}{\partial x_l}\right)\frac{\partial}{\partial x_k}-\sum_{l=1}^{N}\left(\sum_{k=1}^{N}a_{jk}(x)\frac{\partial \beta_l}{\partial x_k}\right)\frac{\partial}{\partial x_l}.\end{aligned}$$

Hence

$$\begin{aligned}\det(Z_1,\dots,Z_{j-1},[T,Z_j],Z_{j+1},\dots,Z_N)&=\det\left(Z_1,\dots,Z_{j-1},\sum_{k=1}^{N}T(a_{jk})\frac{\partial}{\partial x_k},Z_{j+1},\dots,Z_N\right)\\ &\quad-\det\left(Z_1,\dots,Z_{j-1},\sum_{k=1}^{N}a_{jk}\frac{\partial\beta_k}{\partial x_k}\frac{\partial}{\partial x_k},Z_{j+1},\dots,Z_N\right)\end{aligned}$$

$$-\det\left(Z_1, \ldots, Z_{j-1}, \sum_{\substack{l=1\\ l\neq k}}^{N} \alpha_{jk}\frac{\partial \beta_l}{\partial x_k}\frac{\partial}{\partial x_l}, Z_{j+1}, \ldots, Z_N\right)$$

$$=A_j-B_j-C_j.$$

Now $\Sigma_{j=1}^N A_j = T(\det(Z_1, \ldots, Z_N))$, according to the standard formula for differentiating a determinant.

Next, by expanding by minors we see that

$$B_j=\sum_{k=1}^{N}(-1)^{j+k}\alpha_{jk}\frac{\partial \beta_k}{\partial x_k}M_{jk}$$

where M_{jk} is the (j, k)th minor of the matrix $(Z_1, \ldots, Z_N)$. Hence

$$\sum_{j=1}^{N} B_j = \sum_{k=1}^{N}\frac{\partial \beta_k}{\partial x_k}\sum_{j=1}^{N}(-1)^{j+k}\alpha_{jk}M_{jk}$$

$$=(\nabla\cdot T)\det(Z_1, \ldots, Z_N).$$

Similarly,

$$C_j=\sum_{k=1}^{N}(-1)^{j+k}\sum_{\substack{l=1\\ l\neq k}}^{N}\alpha_{jl}\frac{\partial \beta_k}{\partial x_l}M_{jk}.$$

So

$$\sum_{j=1}^{N} C_j = \sum_{k=1}^{N}\sum_{\substack{l=1\\ l\neq k}}^{N}\frac{\partial \beta_k}{\partial x_l}\sum_{j=1}^{N}(-1)^{j+k}\alpha_{jl}M_{jk}$$

$$=\sum_{k=1}^{N}\sum_{\substack{l=1\\ l\neq k}}^{N}\frac{\partial \beta_k}{\partial x_l}\det(Z_1, \ldots, Z_{k-1}, Z_l, Z_{k+1}, \ldots, Z_N)$$

$$=0$$

since each determinant has a repeated row. This proves the lemma.

Remark. The formula is just the Lie derivative of

$$\det(Z_1, \ldots, Z_N)=(Z_1\wedge\ldots\wedge Z_N, dx_1\wedge\ldots\wedge dx_N)$$

with respect to the vector field T.

We can now use Lemma 2.6 to obtain formulas for derivatives of λ_I on the set Ω_I.

LEMMA 2.7. *For every s-tuple $J=(j_1,\dots,j_s)$ of positive integers, there exists $f_J\in A_s^{-d(J)}$ so that*

$$Y_J\lambda_I=f_J\cdot\lambda_I \quad \text{on } \Omega_I.$$

Proof. If $s=1$, Lemma 2.6 shows

$$Y_j\lambda_I=(\nabla\cdot Y_j)\lambda_I+\sum_{k=1}^{N}\det(Y_{i_1},\dots,[Y_j,Y_{i_k}],\dots,Y_{i_N}).$$

On the other hand, Lemma 2.5 shows that

$$[Y_j,Y_{i_k}]=\sum_{l=1}^{N}a^l_{j,i_k}(x)\,Y_{i_l}$$

with

$$a^l_{j,i_k}\in A_1^{d_{i_l}-d_j-d_{i_k}}.$$

Making the substitution, and noting that a determinant with repeated rows is zero, we obtain

$$Y_j\lambda_I=\left(\nabla\cdot Y_j+\sum_{k=1}^{N}a^k_{j,i_k}\right)\lambda_I$$

on Ω_I, which proves the lemma in this case since 1 and hence $\nabla\cdot Y_j$ is an element of $A_1^{-d_j}$. We can now use induction, Leibnitz's rule, and Lemma 2.3 to complete the proof in general.

We also need formulas for derivatives of the determinants λ_J when $J\neq I$.

LEMMA 2.8. *For every N tuple $J=(j_1,\dots,j_N)$ and every s-tuple $K=(k_1,\dots,k_s)$ there exists $f_{K,J}\in A_{N+s}^{d(I)-d(J)-d(K)}$ so that*

$$Y_K\lambda_J=f_{K,J}\lambda_I$$

on the set Ω_I.

Proof. First note that

$$\begin{aligned}\lambda_J&=\det(Y_{j_1},\dots,Y_{j_N})\\&=\det\left(\sum_{l_1=1}^{N}a^{l_1}_{j_1}Y_{i_{l_1}},\dots,\sum_{l_N=1}^{N}a^{l_N}_{j_N}Y_{i_{l_N}}\right).\end{aligned}$$

Expanding this determinant, we see that

$$\lambda_J = f_J \lambda_I$$

with $f_J \in A_N^{d(I)-d(J)}$. We now obtain the lemma by using Leibnitz's rule, and Lemmas 2.4 and 2.7.

§ 2. Estimates on the balls $B_2(x, \delta)$

In § 1, we made no hypothesis about the collections of vector fields $\{Y_j\}$ which guarantees that $\lambda_I(x) \neq 0$ for any $x \in \Omega$ and any N-tuple I. We now return to the basic hypotheses of Chapter 1, and assume that the family $\{Y_j\}$ is finite, say $Y_1, \ldots, Y_q$ and that for each $x \in \Omega$ $Y_1(x), \ldots, Y_q(x)$ span $\mathbf{R}^N$. Then of course for each $x \in \Omega$ there is an N-tuple $I=(i_1, \ldots, i_N)$ so that $\lambda_I(x) \neq 0$.

Our first object is to obtain estimates on elements of the modules A_s^p at a fixed point $x_0 \in \Omega$. We have

LEMMA 2.9. *Let $E \subset\subset \Omega$ be a compact subset. Let $x_0 \in E$ and let I be an N tuple satisfying*

$$|\lambda_I(x_0)|\, \delta^{d(I)} \geq t \sup_J |\lambda_J(x_0)|\, \delta^{d(J)} \tag{16}$$

where $0 < t \leq 1$, $\delta > 0$, and the supremum is over all N tuples. Then:

(a) *If K and L are s- and r-tuples respectively, then*

$$|Y_K a_L^l(x_0)| \leq C t^{-s-1} \delta^{d_{i_l} - d(K) - d(L)}$$

where C depends on K, L, l, I, and the compact set E, but not on t or δ.

(b) *If K and J are s- and N-tuples respectively, then*

$$|Y_K \lambda_J(x_0)| \leq C t^{-s-N} \delta^{d(I)-d(J)-d(K)} |\lambda_I(x_0)|$$

where C depends on K, J, I and E, but not on t or δ.

Proof. According to Lemma 2.2, $a_j^l(x_0) = \lambda_J(x_0)/\lambda_I(x_0)$ where J is obtained from I by replacing i_l by j. But then (16) gives

$$|a_j^l(x_0)| \leq t^{-1} \delta^{d_{i_l} - d_j}.$$

Now Lemma 2.5 shows that

$$a_L^l \in A_1^{d_{i_l}-d(L)},$$

and hence Lemma 2.4 shows that

$$Y_K a_L^l \in A_{s+1}^{d_{i_l}-d(L)-d(K)}.$$

It follows that

$$|Y_K a_L^l(x_0)| \leq C t^{-s-1} \delta^{d_{i_l}-d(K)-d(L)}$$

where the constant C depends on the supremum over E of a finite number of elements of $C^\infty(\Omega)$. This gives (a).

Similarly, Lemma 2.8 shows that $Y_K \lambda_J(x_0) = f_{K,J}(x_0)\,\lambda_I(x_0)$ where $f_{K,J} \in A_{N+s}^{d(I)-d(J)-d(K)}$, and in the same way, this yields (b).

Remark. Since there are only finitely many N-tuples, if we fix an integer n_0, and restrict r and s in Lemma 2.9 to be no bigger than n_0, we can then choose a constant C in Lemma 9 which depends *only* on the compact set E, and is independent of the particular tuples we choose.

In Lemma 2.9 we obtained estimates for various functions at a fixed point x_0. We now want to obtain similar estimates holding on appropriate "balls". Recall from Chapter 1 that we defined a pseudo-distance ϱ_2, which gives rise to a family of "balls" $\{B_2(x,\delta)\}$. In terms of exponential mappings, the balls B_2 are given by

$$B_2(x,\delta) = \left\{ y \in \Omega \mid y = \exp\left(\sum_1^q u_j Y_j\right)(x) \quad \text{with } |u_j| < \delta^{d_j} \right\}.$$

We begin by studying the behaviour of the function $\lambda_I(x)$ on the balls $B_2(x_0,\delta)$ if I satisfies (16).

LEMMA 2.10. *Let $E \subset\subset \Omega$ be a compact set and let $t>0$. There exists $\varepsilon_0(t)>0$ so that if $x_0 \in E$ and if I is an N-tuple satisfying*

$$|\lambda_I(x_0)|\,\delta^{d(I)} \geq t \sup_J |\lambda_J(x_0)|\,\delta^{d(J)} \tag{16}$$

then if $y \in B_2(x_0, \varepsilon_0(t)\,\delta)$

$$|\lambda_I(x_0) - \lambda_I(y)| < \tfrac{1}{2}|\lambda_I(x_0)|.$$

In particular $\lambda_I(y)\neq 0$, *so* $B_2(x_0, \varepsilon_0(t)\,\delta)\subset\Omega_I$.

Proof. Set $F_x(u_1, \ldots, u_q)=\lambda_I(\exp(\Sigma_{j=1}^q u_j Y_j)(x))$. Then F_x is C^∞ on the set $\{u\in\mathbf{R}^N|\ |u|<\delta_0\}$, and F_x also depends smoothly on x. In particular, given E and an integer n, there is a constant A_n so that for $|u|<\delta_0$

$$\left|F_x(u)-\sum_{|\alpha|\leq n}\frac{1}{\alpha!}\frac{\partial^\alpha F_x}{\partial u^\alpha}(0)\,u^\alpha\right|\leq A_n|u|^{n+1}. \tag{17}$$

Now $\inf_{x\in E}\sup_J|\lambda_J(x)|=\eta>0$, where the supremum is over all N-tuples $J=(j_1, \ldots, j_N)$ with $1\leq j_k\leq q$, since the vectors $Y_1, \ldots, Y_q$ are assumed to span $\mathbf{R}^N$ at each point. If we let $n=\sup_{J,K}|d(J)-d(K)|$ where J and K run over all such N-tuples, it follows from (16) that

$$|\lambda_I(x_0)|\geq\eta t\,\delta^n. \tag{18}$$

If we use this choice of n in (17), then if $|u_j|<(\varepsilon_0(t)\,\delta)^{d_j}$

$$A_n|u|^{n+1}<\tfrac{1}{4}|\lambda_I(x_0)| \tag{19}$$

if $\varepsilon_0(t)$ is small enough, depending only on the compact set E. Thus in order to prove the lemma, it suffices to show

$$\left|\sum_{1\leq|\alpha|\leq n}\frac{1}{\alpha!}u^\alpha\frac{\partial^\alpha F_{x_0}}{\partial u^\alpha}(0)\right|\leq\tfrac{1}{4}|\lambda_I(x_0)|. \tag{20}$$

But (see the appendix, Proposition 4.2)

$$\sum_{|\alpha|=s}\frac{1}{\alpha!}u^\alpha\frac{\partial^\alpha F_{x_0}}{\partial u^\alpha}(0)=\frac{1}{s!}\left(\sum_{j=1}^q u_j Y_j\right)^s\lambda_I(x_0). \tag{21}$$

A typical term from the right hand side of (21) is

$$\frac{1}{s!}u_{j_1}\ldots u_{j_s}Y_{j_1}\ldots Y_{j_s}\lambda_I(x_0).$$

If $|u_j|<(\varepsilon_0(t)\,\delta)^{d_j}$, Lemma 2.9 (b) gives the estimate

$$C\frac{1}{s!}\varepsilon_0(t)^{d(J)_I-s-N}|\lambda_I(x_0)|.$$

Thus we can choose $\varepsilon_0(t)$ sufficiently small, depending only on the compact set E, so that (20) is satisfied, and the lemma is proved.

A similar argument, again using Lemma 2.9(b), shows that we can obtain estimates for $|\lambda_J(y)|$ if $y \in B_2(x_0, \varepsilon_0(t)\delta)$ if $J \neq I$. We do not repeat the proof.

LEMMA 2.11. *There is a constant C depending only on the compact set $E \subset\subset \Omega$ so that if $\varepsilon_0(t)$ is sufficiently small, if (16) is satisfied, and if $y \in B_2(x_0, \varepsilon_0(t)\delta)$ then*

$$|\lambda_J(y)| \leqslant Ct^{-N}\delta^{d(I)-d(J)}|\lambda_I(x_0)|.$$

If we now use Lemma 2.2, and the estimates of Lemmas 2.10 and 2.11 we see that if (16) is satisfied, if $\varepsilon_0(t)$ is sufficiently small, and if $y \in B_2(x_0, \varepsilon_0(t)\delta)$ then

$$|a_j^l(y)| \leqslant Ct^{-N}\delta^{d_{i_l}-d_j}$$

where C depends only on the compact set E. But then we can repeat the proof of Lemma 2.9 to obtain estimates on $B_2(x_0, \varepsilon_0(t)\delta)$.

THEOREM 6. *Let $E \subset\subset \Omega$ be a compact set, $t>0$, and n_0 a fixed integer. There are constants C, and $\varepsilon(t)$ with the following properties: let $x_0 \in E$ and let I be an N-tuple satisfying*

$$|\lambda_I(x_0)|\,\delta^{d(I)} \geqslant t \max_J |\lambda_J(x_0)|\,\delta^{d(J)}. \tag{16}$$

Then if $y \in B_2(x_0, \varepsilon(t)\delta)$

(a) *If K and L are s- and r-tuples with s and r at most n_0, then*

$$|Y_K a_L^l(y)| \leqslant Ct^{-s-1}\delta^{d_{i_l}-d(K)-d(L)}.$$

(b) *If K and J are s- and N-tuples with $s \leqslant n_0$ then*

$$|Y_K \lambda_J(y)| \leqslant Ct^{-N-s}\delta^{d(I)-d(J)-d(K)}|\lambda_I(x_0)|.$$

§3. The Main structure theorem

Before proceeding, we introduce a simplification of notation. If $x \in E$ and $I=(i_1, \ldots, i_N)$ are fixed, we shall relabel the vector fields $\{Y_j\}$ by setting $U_j = Y_{i_j}$, $1 \leqslant j \leqslant N$, and by letting V_j, $1 \leqslant j \leqslant q-N$, be some enumeration of the remaining vector fields. If $u=(u_1, \ldots, u_N) \in \mathbf{R}^N$ and $v=(v_1, \ldots, v_{q-N}) \in \mathbf{R}^{q-N}$, we let

$$u \cdot U + v \cdot V = \sum_{j=1}^{N} u_j U_j + \sum_{j=1}^{q-N} v_j V_j \tag{22}$$

and we set

$$\Phi_v(u) = \exp(u \cdot U + v \cdot V)(x). \tag{23}$$

For $v \in \mathbf{R}^{q-N}$, we let

$$z = \exp(v \cdot V)(x)$$

and we introduce one more family of balls

$$B_I(x, z, \delta) = \{y \in \Omega \,|\, y = \exp(u \cdot U + v \cdot V)(x), \text{ with } |u_j| < \delta^{d(U_j)}\}. \tag{24}$$

Thus $B_I(x, z, \delta)$ is exactly the image, under the map Φ_v of the box $\{u \in \mathbf{R}^N \,|\, |u_j| < \delta^{d(U_j)}\} = Q(\delta)$.

We can now state our main result on the structure of the balls $\{B(x, \delta)\}$.

THEOREM 7. *Let $E \subset\subset \Omega$ be compact. There exist constants $0 < \eta_2 < \eta_1 < 1$ so that if $x \in E$ and $\delta > 0$ there exists an N-tuple $I = (i_1, \dots, i_N)$ with the following properties:*

(1) $$|\lambda_I(x)|\, \delta^{d(I)} \geq \eta_2 \max_J |\lambda_J(x)|\, \delta^{d(J)}.$$

(2) *If $|v_j| < \eta_2 \delta^{d(V_j)}$, $1 \leq j \leq q-N$, Φ_v is one-to-one on the box $Q(\eta_1 \delta)$.*

(3) *If $|v_j| < \eta_2 \delta^{d(V_j)}$, $1 \leq j \leq q-N$, Φ_v is nonsingular on the box $Q(\eta_1 \delta)$, and if $|J\Phi_v|$ denotes the Jacobian, then on $Q(\eta_1 \delta)$,*

$$\tfrac{1}{4}|\lambda_I(x)| \leq |J\Phi_v| \leq 4|\lambda_I(x)|.$$

(4) *If $|v_j| < \eta_2 \delta^{d(V_j)}$, $1 \leq j \leq q-N$, and $z = \exp(v \cdot V)(x)$*

$$B(x, \eta_2 \delta) \subset B_I(x, z, \eta_1 \delta) \subset B_2(x, \delta) \subset B(x, \delta).$$

Our first goal is to establish (3). We show that if $x \in E$, and I is an N-tuple with

$$|\lambda_I(x)|\, \delta^{d(I)} \geq t \max_J |\lambda_J(x)|\, \delta^{d(J)} \tag{16}$$

then Φ_v is nonsingular, and hence locally one-to-one on the box

$$\{u \in \mathbf{R}^N \,|\, |u_j| < (\varepsilon(t)\,\delta)^{d(U_j)}\}$$

provided that $|v_j| < (\varepsilon(t)\,\delta)^{d(V_j)}$, $1 \leq j \leq q-N$, and provided that $\varepsilon(t)$ is sufficiently small, depending only on the compact set $E \subset\subset \Omega$ and t. To do this, we must compute the Jacobian of Φ_v.

More generally, let $W_1, \ldots, W_q$ be any q vector fields on Ω, and set

$$\theta_x(s) = \exp\left(\sum_{j=1}^{q} s_j W_j\right)(x). \tag{25}$$

LEMMA 2.12. *Let $E \subset\subset \Omega$ be compact. Then for any integer n there are constants $\delta_0, \alpha_2, \ldots, \alpha_n$, and C_n so that if $x_0 \in E$*

$$\left| d\theta_{x_0}\left(\frac{\partial}{\partial s_j}\right)(\theta_{x_0}(s)) - \left\{ W_j + \sum_{k=2}^{n} \alpha_k [s \cdot W, [\ldots [s \cdot W, W_j] \ldots]] \right\} \right| \leq C_n |s|^{n+1} \quad \text{if } |s| < \delta_0 \tag{26}$$

where the commutators on the left hand side of (26) *are of length k.*

Proof. If we could write

$$\exp(s \cdot W + tW_j)(x) = \exp(tZ_j) \circ \exp(s \cdot W)(x)$$

the definition of the exponential map would give

$$d\theta_x\left(\frac{\partial}{\partial s_j}\right) = Z_j.$$

To compute Z_j, we use the Campbell-Hausdorff formula, as in Proposition 4.3 of the appendix. For any positive integer n, and any $x \in E$ we have

$$|\exp(s \cdot W + tW_j) \circ \exp(-s \cdot W)(x) - \exp R_n(x)| \leq C_n(|s|^{n+1} t + t^2)$$

where $R_n = tW_j + \Sigma_{k=2}^{n} \alpha_k [s \cdot W, [s \cdot W, \ldots [s \cdot W, W_j] \ldots]]$. We now let $x = \exp(s \cdot W) x_0)$, and the lemma is proved.

Now let $W_i = U_i$ for $i = 1, \ldots, N$, and $W_i = V_{i-N}$ for $i = N+1, \ldots, q$. We can then prove

LEMMA 2.13. *Let $E \subset\subset \Omega$ be compact, let $\varkappa > 0$, and let $t > 0$. There exists $\varepsilon(t) > 0$ so that if $x_0 \in E$ and if I is an N-tuple satisfying*

$$|\lambda_I(x_0)|\, \delta^{d(I)} \geq t \sup_J |\lambda_J(x_0)|\, \delta^{d(J)} \tag{16}$$

then if $|u_j| < (\varepsilon(t)\,\delta)^{d(U_j)}$, $1 \leq j \leq N$, and $|v_j| < (\varepsilon(t)\,\delta)^{d(V_j)}$, $1 \leq j \leq q-N$, then on $B(x_0, \varepsilon(t)\,\delta)$

$$d\Phi_v\left(\frac{\partial}{\partial u_j}\right)(\Phi_v(u)) - U_j = \sum_{l=1}^{N} b_{jl} U_l$$

where $|b_{jl}|\leqslant\varkappa\,\delta^{d(U_l)-d(U_j)}$.

Proof. According to Lemma 2.12, for any integer n

$$d\Phi_v\left(\frac{\partial}{\partial u_j}\right)(\Phi_v(u))-U_j=\sum_{k=2}^{n}\alpha_k[u\cdot U+v\cdot V,\,\dots,\,[u\cdot U+v\cdot V,\,U_j]\dots]+O((|u|+|v|)^{n+1}).$$

Arguing as in Lemma 2.10, we can choose n so large that the term $O((|u|+|v|)^{n+1})$ has the right form if $\varepsilon(t)$ is small enough. On the other hand, we can use Theorem 6 (a) to estimate each term in the finite sum, and again the sum has the right form if $\varepsilon(t)$ is small enough.

We can now complete the proof of part (3) of Theorem 7.

LEMMA 2.14. *Let $E\subset\subset\Omega$ be a compact subset and $t>0$. There exists $\varepsilon(t)>0$ so that if $x_0\in E$, if I is an N-tuple satisfying*

$$|\lambda_I(x_0)|\,\delta^{d(I)}\geqslant t\sup_J|\lambda_J(x_0)|\,\delta^{d(J)} \tag{16}$$

if $|u_j|<(\varepsilon(t)\,\delta)^{d(U_j)}$, $1\leqslant j\leqslant N$, *and* $|v_j|<(\varepsilon(t)\,\delta)^{d(V_j)}$, $1\leqslant j\leqslant q-N$, *then*

$$\tfrac{1}{4}|\lambda_I(x_0)|\leqslant|J\Phi_v(u)|\leqslant4|\lambda_I(x_0)|.$$

Proof.

$$J\Phi_v=\det\left(d\Phi_v\left(\frac{\partial}{\partial u_1}\right),\,\dots,\,d\Phi_v\left(\frac{\partial}{\partial u_N}\right)\right).$$

However

$$|\det(U_1,\dots,U_N)(\Phi_v(u))|=|\lambda_I(\Phi_v(u))|$$

with $\frac{1}{2}|\lambda_I(x_0)|\leqslant|\lambda_I(\Phi_v(u))|\leqslant2|\lambda_I(x_0)|$ if $\varepsilon(t)$ is small by Lemma 2.10. But now Lemma 2.14 follows from Lemma 2.13 if we choose $\varkappa$ small enough.

If $|u_j|<(\varepsilon(t)\,\delta)^{d(U_j)}$ and if $y=\Phi_v(u)$, we have two N-tuples at y: the set $(U_1,\dots,U_N)$, and the set $d\Phi_v(\partial/\partial u_1),\dots,d\Phi_v(\partial/\partial u_N)$. If we let $Z_j=d\Phi_v(\partial/\partial u_j)$, then Lemma 2.13 says

$$Z_j=U_j+\sum_{l=1}^{N}b_{jl}\,U_l=\sum_{l=1}^{N}(\delta_{jl}+b_{jl})\,U_l \tag{27}$$

where $|b_{jl}|<\varkappa\,\delta^{d(U_l)-d(U_j)}$. We can, of course, solve for the U_l in terms of the Z_j. Let M_{jl} be the (j,l)th minor of the matrix $\{\delta_{jl}+b_{jl}\}$. If we set

$$C_{l,k}=(-1)^{l+k}\frac{\det M_{kl}}{\det(\delta_{ij}+b_{ij})}$$

then

$$U_l=\sum_{k=1}^{N}C_{lk}Z_k. \tag{28}$$

But it is easy to see that if $\varkappa$ is sufficiently small, uniformly in δ, then $|\det(\delta_{ij}+b_{ij})|\geq\frac{1}{2}$, while $|\det M_{kl}|\leq C\,\delta^{d(U_k)-d(U_l)}$. Thus

$$|C_{lk}|\leq C\,\delta^{d(U_k)-d(U_l)}. \tag{29}$$

We can reinterpret equation (29). Since $J\Phi_V(u)\neq 0$, there is an open neighborhood of $y=\Phi_V(u)$ on which Φ_V has an inverse map $\Psi=(\psi_1,\ldots,\psi_N)$, so that locally

$$\psi_j(\Phi_V(u))=u_j.$$

We can regard $\psi_1,\ldots,\psi_N$ as coordinate functions near y, and relative to these coordinates, the coefficient C_{lk} is just the kth coordinate of U_l; i.e.

$$C_{lk}=U_l(\psi_k). \tag{30}$$

Thus we have proved

LEMMA 2.15. *Let $E\subset\Omega$ be compact, and let $t>0$. There are constants C and $\varepsilon(t)$ so that if $x_0\in E$, if I is an N-tuple satisfying*

$$|\lambda_I(x_0)|\,\delta^{d(I)}\geq t\max_J|\lambda_J(x_0)|\,\delta^{d(J)}, \tag{16}$$

if $|u_j|<(\varepsilon(t)\,\delta)^{d(U_j)}$, $1\leq j\leq N$, if $|v_j|<(\varepsilon(t)\,\delta)^{d(V_j)}$, $1\leq j\leq q-N$, if $y=\exp(u\cdot U+v\cdot V)(x_0)$, and if $\Psi=(\psi_1,\ldots,\psi_N)$ is locally the inverse map to Φ_V then

$$|U_l(\psi_k)|\leq C\,\delta^{d(U_k)-d(U_l)}.$$

We now turn to the proof of part (4) of Theorem 7. Let $E\subset\Omega$ be compact, let $x\in E$, let $\delta>0$, and suppose I is an N-tuple such that

$$|\lambda_I(x)|\,\delta^{d(I)}\geq t\max_J|\lambda_J(x)|\,\delta^{d(J)}.$$

It is obvious from the definitions that if $|v_j|<\delta^{d(V_j)}$ for $1\leq j\leq q-N$ then we have the inclusions

$$B_I(x, z, \delta) \subset B_2(x, \delta) \subset B(x, \delta)$$

where $z=\exp(v\cdot V)(x)$. Thus the difficulty in part (4) is proving inclusions of the form $B(x, \eta_2\delta)\subset B_I(x, z, \eta_1\delta)$. We begin with the following result.

LEMMA 2.16. *There exists $\eta>0$ so that if $x_0\in E$, if I is an N-tuple for which*

$$|\lambda_I(x_0)|\,\delta^{d(I)} \geq t \sup_J |\lambda_J(x_0)|\,\delta^{d(J)} \tag{16}$$

if $|v_j|<(\varepsilon(t)\,\delta)^{d(V_j)}$, and if $z=\exp(v\cdot V)(x_0)$ then

$$B(z, \eta\varepsilon(t)\,\delta) \subset B_I(x_0, z, \varepsilon(t)\,\delta).$$

Proof. Let $y\in B(z, \eta\varepsilon(t)\,\delta)$. Then there is an absolutely continuous map $\varphi\colon [0, 1]\to\Omega$ with $\varphi(0)=z$, $\varphi(1)=y$ and

$$\varphi'(t) = \sum_{j=1}^{q} b_j(t)\, Y_j(\varphi(t)) \tag{31}$$

with $|b_j(t)|<(\eta\varepsilon(t)\,\delta)^{d_j}$. We can also assume that the map φ is one to one.

Let $\mathscr{S}$ be the set of numbers $s_0\in[0, 1]$ such that there exists an absolutely continuous mapping $\theta\colon[0, s_0]\to\mathbf{R}^N$ such that $|\theta_j(s)|\leq(\frac{1}{2}\varepsilon(t)\,\delta)^{d(v_j)}$ and $\varphi(s)=\exp(\Sigma_{j=1}^N \theta_j(s)\, U_j+v\cdot V)(x_0)$, $0\leq s\leq s_0$.

Since the mapping

$$(u_1, \ldots, u_N) \mapsto \exp(u\cdot U+v\cdot V)(x_0)$$

is locally one to one on $\{u\in\mathbf{R}^N |\, |u_j|<(\varepsilon(t)\,\delta)^{d(U_j)}\}$ it follows that the components θ_j are unique. In fact if we had two such θ's, say θ^1 and θ^2, the set where $\theta^1=\theta^2$ would be open and closed, and would contain a small neighborhood of the origin. We let

$$\bar{s} = \sup\{s_0\in\mathscr{S}\}.$$

We want to show that if η is sufficiently small, $\bar{s}=1$, for then

$$y = \varphi(1) = \exp\left(\sum_{j=1}^{N} \theta_j(1)\, U_j+v\cdot V\right)(x_0)$$

with $|\theta_j(1)|<(\varepsilon(t)\,\delta)^{d(U_j)}$ and so $y\in B_I(x_0, z, \varepsilon(t)\,\delta)$.

The mapping $\Phi_v(u_1, \ldots, u_N)=\exp(u\cdot U+v\cdot V)(x_0)$ is locally one to one, and since the map φ and θ are one to one on $[0, \bar{s}]$, and

$$\varphi(s)=\Phi_v(\theta(s))$$

it follows that Φ_v is actually globally one-to-one on some small neighborhood of the image $\theta[0, \bar{s}]$. Thus we can think of the components of the inverse map $(\psi_1, \ldots, \psi_N)$ as being well defined functions in some neighborhood of $\theta([0, \bar{s}])$.

If $\bar{s}<1$, for some j_0 we must have $\psi_{j_0}(\bar{s})=(\frac{1}{2}\varepsilon(t)\,\delta)^{d(U_{j_0})}$. On the other hand, for any j_0 we have

$$\begin{aligned}\psi_{j_0}(\bar{s}) &= \psi_{j_0}(\bar{s})-\psi_{j_0}(0)\\ &=\int_0^{\bar{s}} \frac{d}{ds}\psi_{j_0}(s)\,ds\\ &=\int_0^{\bar{s}}\sum_{j=1}^{q} b_j(s)\,Y_j(\varphi(s))\,(\psi_{j_0})\,ds\\ &=\int_0^{\bar{s}}\sum_{j=1}^{q}\sum_{l=1}^{N} b_j(s)\,a_j^l(\varphi(s))\,U_l(\varphi(s))\,(\psi_{j_0})\,ds.\end{aligned}$$

But $|U_l(\psi_{j_0})|\le C\,\delta^{d(U_{j_0})-d(U_l)}$ by Lemma 2.15. Thus

$$\begin{aligned}|\psi_{j_0}(\bar{s})&\le(\eta\varepsilon(t)\,\delta)^{d_j}Ct^{-1}\delta^{d(U_l)-d_j}C\,\delta^{d(U_{j_0})-d(U_l)}\\ &<(\tfrac{1}{2}\varepsilon(t)\,\delta)^{d(U_{j_0})}\end{aligned}$$

if η is small enough. Thus $\bar{s}=1$, and the proof is complete.

Note that if we choose $v=0$, we have shown that

$$B(x_0, \eta\varepsilon(t)\,\delta)\subset B_I(x_0, \varepsilon(t)\,\delta). \tag{32}$$

We can easily take care of the case $v\neq 0$ if v is sufficiently small. We have, with the notation of the previous lemma

LEMMA 2.17. *Suppose* $|v_j|<(\frac{1}{3}\eta\varepsilon(t)\,\delta)^{d(V_j)}$. *Then*

$$B\left(x_0, \frac{\eta}{3}\varepsilon(t)\,\delta\right)\subset B_I(x_0, z, \varepsilon(t)\,\delta).$$

Proof. Let $y \in B(x_0, \frac{1}{3}\eta\varepsilon(t)\delta)$. Thus $\varrho(x_0, y) < \frac{1}{3}\eta\varepsilon(t)\delta$. But $\varrho(x_0, z) < \frac{1}{3}\eta\varepsilon(t)\delta$ by hypothesis, so $\varrho(y, z) < \frac{2}{3}\eta\varepsilon(t)\delta$. Hence $B(x_0, \frac{1}{3}\eta\varepsilon(t)\delta) \subset B(z, \frac{2}{3}\eta\varepsilon(t)\delta)$.

Finally, we turn to the proof of part (2) of Theorem 7. Let $E \subset\subset \Omega$ be compact and let $x_0 \in E$. Let $I_0 = I_0(x)$ be an N-tuple such that $d(I_0)$ is minimal among all N-tuples I with $\lambda_I(x_0) \neq 0$, and such that

$$|\lambda_{I_0}(x_0)| = \max_{d(I)=d(I_0)} |\lambda_I(x_0)|. \tag{33}$$

Then there exists δ_0 depending on x_0 such that

$$|\lambda_{I_0}(x_0)|\,\delta^{d(I_0)} \geq |\lambda_I(x_0)|\,\delta^{d(I)} \tag{34}$$

for all δ, $0 < \delta \leq \delta_0$, and all N-tuples I. Also, by choosing a smaller δ_0 if necessary, we can find an open set W in Ω containing x_0 so that the mapping

$$(u_1, \dots, u_N) \mapsto \Phi_v(u_1, \dots, u_N) = \exp(u \cdot U + v \cdot V)(x)$$

is globally one to one on $|u| < \delta_0$ for all x in W, $|v| < \delta_0$. This is true since the Jacobian of the exponential map is the identity at the origin. (See the appendix.) We may also assume

$$|\lambda_{I_0}(x)|\,\delta_0^{d(I_0)} > \tfrac{1}{2}|\lambda_I(x)|\,\delta_0^{d(I)} \tag{35}$$

for all N-tuples I, and all $x \in W$.

Let K be a compact subset of W containing x_0. For $x \in K$, we can choose a sequence of N-tuples $I_1, \dots, I_n$, and real numbers $\delta_0 > \delta_1 > \dots > \delta_n > 0$ so that for

$$\begin{gathered} \delta_{j+1} \leq \delta \leq \delta_j, \quad 0 \leq j \leq n-1 \\ |\lambda_{I_j}(x)|\,\delta^{d(I_j)} \geq \frac{1}{2} \sup_J |\lambda_J(x)|\,\delta^{d(J)} \end{gathered} \tag{36}$$

while for

$$\begin{gathered} 0 < \delta \leq \delta_n \\ |\lambda_{I_n}(x)|\,\delta^{d(I_n)} \geq \frac{1}{2} \sup_J |\lambda_J(x)|\,\delta^{d(J)}. \end{gathered} \tag{37}$$

We may clearly assume $d(I_{j+1}) < d(I_j)$. In particular, no N-tuple occurs twice, and n is at most the total number of allowable N-tuples. The choice of the particular N-tuple of course may depend on x. According to Lemma 2.17 there is a function $\varepsilon(t)$ so that if for some $x\in E$, $0<\delta<\delta_0$, and N-tuple I, $|\lambda_I(x)|\,\delta^{d(I)}\geq t\sup_J|\lambda_J(x)|\,\delta^{d(J)}$, then if $|v_j|<(\varepsilon(t)\,\delta)^{d(V_j)}$, $B(x,\varepsilon(t)\,\delta)\subset B_I(x,v,\delta)$. We now prove:

LEMMA 2.18. *Suppose for some* δ

$$|\lambda_{I_j}(x)|\,\delta^{d(I_j)}\geq\frac{1}{2}\sup_J|\lambda_J(x)|\,\delta^{d(J)}$$

for $j=0,1$. *Put* $\eta_1=\varepsilon(\frac{1}{2})$ *and* $\eta_2=\varepsilon(\frac{1}{2}\eta_1^{n_0})\,\eta_1$ *where* $n_0=\sup_{I,J}|d(I)-d(J)|$. *Then if* $|v_j|<(\eta_2\,\delta)^{d(V_j)}$

$$B_{I_1}(x,v,\eta_2\,\delta)\subset B_{I_0}(x,\eta_1\,\delta)\subset B_{I_1}(x,v,\delta).$$

Proof. $B_{I_1}(x,v,\delta)\supset B(x,\varepsilon(\frac{1}{2})\,\delta)\supset B_{I_0}(x,\varepsilon(\frac{1}{2})\,\delta)$ by the theorem, and the definition. On the other hand

$$\begin{aligned}|\lambda_{I_0}(x)|\,(\eta_1\,\delta)^{d(I_0)}&\geq\tfrac{1}{2}\eta_1^{d(I_0)-d(J)}|\lambda_J(x)|\,(\eta_1\,\delta)^{d(J)}\\&\geq\tfrac{1}{2}\eta_1^{n_0}|\lambda_J(x)|\,(\eta_1\,\delta)^{d(J)}\end{aligned}$$

for any N-tuple J. Hence we can apply Lemma 2.17 again, with $t=\frac{1}{2}\eta_1^{n_0}$ and $\delta'=\eta_1\,\delta$. We get

$$B_{I_0}(x,\eta_1\,\delta)\supset B(x,\varepsilon(\tfrac{1}{2}\eta_1^{n_0})\,\eta_1\,\delta)\supset B_{I_1}(x,v,\eta_2\,\delta)$$

which proves the lemma.

Now we know that the mapping

$$(u_1,\ldots,u_N)\mapsto\exp(u\cdot U+v\cdot V)(x)$$

is globally one-to-one if $x\in K$, $|u|<\delta_0$, $|v|<\delta_0$. In particular, it follows that the image of any simply connected set is simply connected. Let $\Phi_v^{(1)}$ be the mapping (23) associated to the N-tuple I_1. If $\Phi_v^{(1)}$ were not globally one-to-one on $|u_i|<(\eta_2\,\delta)^{d(U_i)}$, there would be a line segment L in the box

$$\{u\in\mathbf{R}^N\,|\,|u_j|<(\eta_2\,\delta)^{d(U_j)}\}$$

which $\Phi_v^{(1)}$ maps to a closed curve in $B_{I_1}(x,z,\eta_2\,\delta)$, where $z=\exp(v\cdot V)(x)$. However, this curve can be deformed to a point in $B_{I_0}(x,\eta_1\,\delta)$ and hence it can be deformed to a point in $B_{I_1}(x,z,\delta)$, which is impossible. Thus $\Phi_v^{(1)}$ is globally one-to-one.

By repeating this argument n times for successive series of N-tuples I_{j+1} and I_j, we prove

LEMMA 2.19. *Let $E \subset\subset \Omega$ be compact. There exist $0<\eta_2<\eta_1<1$ so that if $x \in E$ and $0<\delta<\delta_0$ there exists an N-tuple I so that*

$$|\lambda_I(x)|\,\delta^{d(I)} \geq \frac{1}{2}\max_J |\lambda_J(x)|\,\delta^{d(J)}$$

and so that if $|v_j|<(\eta_2\,\delta)^{d(V_j)}$, $1 \leq j \leq q-N$, the mapping

$$(u_1, \ldots, u_N) \mapsto \exp(u\cdot U+v\cdot V)(x)$$

is globally one-to-one for $|u_j|<(\eta_1\,\delta)^{d(U_j)}$.

We have now completely proved Theorem 7.

§ 4. Proofs of Theorems 1, 2, and 3

As immediate corollaries, we can now prove Theorems 1, 2 and 3 of Chapter 1. First, since $B_I(x, z, \eta_1\,\delta)$ is the image under the one to one mapping Φ_v of the box $Q(\eta_1\,\delta)$, and since the Jacobian of this mapping is bounded between two constant multiples of $\lambda_I(x)$, it follows that

$$\begin{aligned} |B_I(x, z, \eta_1\,\delta)| &\approx |\lambda_I(x)|\,|Q(\eta_1\,\delta)| \\ &\approx |\lambda_I(x)|\,\delta^{d(I)}. \end{aligned}$$

Moreover, since $B(x, \eta_2\,\delta) \subset B_I(x, z, \eta_1\,\delta) \subset B(x, \delta)$ it follows that

$$|B(x, \delta)| \approx \sum_I |\lambda_I(x)|\,\delta^{d(I)}$$

which proves Theorem 1.

Theorems 2 and 3 are immediate from part (4) of Theorem 7 since for $x \in E$ and all small δ there is an I so that

$$B(x, \eta_2\,\delta) \subset B_I(x, 0, \eta_1\,\delta) \subset B_2(x, \delta) \subset B(x, \delta).$$

§ 5. Proof of Theorem 5

In this section, we prove that ϱ and ϱ_4 are locally equivalent. It is clear from the definitions that $\varrho \leq \varrho_4$. Thus we must show that for any z in Ω there is an open set U

containing z so that there exists a constant C with

$$\varrho_4(x,y) \leq C\varrho(x,y)$$

for any x, y in U.

Our proof is a modification of an argument of Hörmander [H]. The main work is to establish the following result:

LEMMA 2.20. *Let $w \in \Omega$ Then w has a neighborhood U so that if x_1 and x_∞ are in U with $\varrho(x_1, x_\infty) < \varepsilon$, then the following two conclusions hold:*

(a) *There exists $x_2 \in U$ with $\varrho_4(x_1, x_2) < C\varepsilon$, and $\varrho(x_2, x_\infty) < C\varepsilon^{1+1/m}$.*

(b) *Given $y \in U$ there is a number $\eta(y) > 0$ so that if $|z-y| < \eta(y)$, $\varrho_4(y,z) < C|z-y|^{1/m}$.*

We first show that Lemma 2.20 implies that ϱ and ϱ_4 are locally equivalent. Given $w \in \Omega$ choose U a neighborhood so small that the preceding lemmas apply. Let $x = x_1$, and y be in U with $\varrho(x,y) = \delta$. We apply Lemma 2.20 with $x_1 = x$, $x_\infty = y$ and obtain a point x_2 with

$$\varrho_4(x_1, x_2) < C\delta \quad \text{and} \quad \varrho(x_2, y) < C\delta^{1+1/m} < \tfrac{1}{2}\delta$$

if $C\delta^{1/m} < \frac{1}{2}$. We can then apply Lemma 2.20 again with $\varepsilon = \frac{1}{2}\delta$ to obtain x_3 so that $\varrho(x_2, x_3) < \frac{1}{2}C\delta$, $\varrho(x_3, y) < \frac{1}{4}\delta$. In general, given $x = x_1, x_2, \dots, x_j$ we can find x_{j+1} so that $\varrho_4(x_j, x_{j+1}) < C(\delta/2^{j-1})$ and $\varrho(x_{j+1}, y) < \delta/2^j$. Moreover ϱ_4 satisfies the triangle inequality so $\varrho_4(x, x_j) < C\delta$.

By part (b) of Lemma 2.20 we see that if j is sufficiently large, $\varrho_4(x_j, y) < \delta$. Using the triangle inequality again for ϱ_4 completes the proof.

The main step in the proof of Lemma 2.20 is the following generalization of the Campbell-Hausdorff formula. Let $S_1, \dots, S_l$ be l vector fields. We define

$$C_1(a, S_1) = e^{aS_1}$$

$$C_2(a, S_1, S_2) = e^{aS_1} e^{aS_2} e^{-aS_1} e^{-aS_2},$$

and in general

$$C_l(a, S_1, \dots, S_l) = e^{aS_1} C_{l-1}(a, S_2, \dots, S_{l-1})\, e^{-aS_1} (C_{l-1}(a, S_2 \dots S_l))^{-1}.$$

Note that when C_l or C_l^{-1} is written out as a composition of mappings $\exp \pm aS_j$ there are $n_l = 3\cdot 2^{l-1}-2$ such terms. According to the Campbell-Hausdorff theorem, if $l=2$,

$$C_2(a, S_1, S_2) = \exp(a^2[S_1, S_2]+R)$$
$$R = \sum_{j=3}^{\infty} a^j S_j$$

where S_j is a commutator of length j. The above formula implies that if one replaces R by $R_k = \Sigma_{3\le j\le k}\, a^j S_j$

$$|C_2(a, S_1, S_2)x - \exp(a^2[S_1, S_2]+R_k)x| \le c|a|^{k+1}.$$

(See the appendix.) By induction on l we obtain:

LEMMA 2.21. *$C_l(a, S_1, S_2, \ldots, S_l) = \exp\{a^l[S_1, [S_2, [\ldots S_l] \ldots] + R\}$*

$$R = \sum_{k>l} c_k a^k \times \textit{commutators of length } k.$$

Here c_k are constants independent of a, $S_1, \ldots, S_l$.

We shall now use Lemma 2.21 to prove the second part of Lemma 2.20. Let $T_1, \ldots, T_N$ be commutators of the $X_0, \ldots, X_p$ spanning the tangent space of Ω at y. Let $d_i = \deg T_i$.

If

$$T_i = \left[X_{j_1}^{(1)}, \left[X_{j_2}^{(1)}, \ldots \left[X_{j_{d_i-1}}^{(1)}, X_{j_{d_i}}^{(1)}\right], \ldots\right]\right.,$$

set

$$C_i(t) = C_{d_i}(t^{1/d_i}, X_{j_1}^{(1)}, \ldots, X_{j_{d_i}}^{(1)})\, y. \tag{38}$$

Then by Lemma 2.21 and Proposition 1.1 $C_i(t)$ is a C^1 map of an interval of t's containing the origin onto a curve in Ω through y, and

$$\left.\frac{dC_i(t)}{dt}\right|_{t=0} \text{ is } T_i. \tag{39}$$

Now define

$$C(t_1, \ldots, t_N)\, y = C_1(t_1)\, C_2(t_2) \ldots C_N(t_N)\, y.$$

Then from (39) we see that the Jacobian of C is non-singular at $t=0$. So C is locally a C^1 map from a neighborhood of the origin in $\mathbf{R}^N$ onto a neighborhood of y in Ω. In view of (38) and the definition of ϱ_4, the second part of Lemma 2.20 is proved. To prove the first part of Lemma 2.20 we need a further lemma. We let η_l=dimension of the space of lth order commutators and set

$$\eta = \sum_{l=1}^{m} \eta_l n_l.$$

(n_l recall was the number of terms $e^{\pm aS_j}$ in $C_l(a, S_1, ..., S_l)$.) η is some fixed but unimportant number.

LEMMA 2.22. *Suppose $a^{(1)}, a^{(2)}, ..., a^{(m)}$ are given constant vectors with $|a^{(j)}|<\delta^j$, $j=1, ..., m$. Then there exists constants c_μ, $\mu=1, ..., \eta$ with $|c_\mu|\leqslant c\delta$, so that*

$$\left| \exp(a^{(1)}\cdot X^{(1)}+...+a^{(m)}\cdot X^{(m)})x - \prod_{\mu=1}^{\eta} \exp(c_\mu X^{(1)}_{j(\mu)})x \right| \leqslant C\delta^{m+1},$$

Lemma 2.22 is implied by Lemma 2.21 as follows. First we take care of $a^{(1)}\cdot X^{(1)}$ by writing a product of $e^{a_0^{(1)}X_0} e^{a_1^{(1)}X_1} ... e^{a_p^{(1)}X_p}$. That is

$$e^{a^{(1)}\cdot X^{(1)}+R} = e^{a_0^{(1)}X_0} e^{a_1^{(1)}X_1} ... e^{a_p^{(1)}X_p}$$

where $R=\Sigma_{k\geqslant 2}\, \mathcal{O}(\delta^k)\times$commutators of length k. Next by using Lemma 2.21 when $l=2$, we can match up commutators of second order giving

$$e^{a^{(1)}\cdot X^{(1)}+a^{(2)}\cdot X^{(2)}+R} = e^{a_0^{(1)}X_0} ... e^{a_p^{(1)}X_p} e^{b_1 X_{j_1}} ... e^{b_{n_2\eta_2} X_{j_{n_2\eta_2}}}$$

where $R=\Sigma_{k\geqslant 3}\, \mathcal{O}(\delta^k)\times$commutators of length k, etc. This proves Lemma 2.22.

We now show that Lemma 2.22 implies the first part of Lemma 2.20: if $\varrho(x_1, x_\infty)<\varepsilon$, $\varrho_2(x_1, x_\infty)<C\varepsilon$. Thus there are $a^{(j)}$, $|a^{(j)}|<(C\varepsilon)^j$ such that

$$x_\infty = \exp(a^{(1)}\cdot X^{(1)}+...+a^{(m)}\cdot X^{(m)})x_1.$$

By Lemma 2.22 we can find x_2 so that $\varrho_4(x_1, x_2)<C\varepsilon$ and $|x_2-x_\infty|\leqslant c\varepsilon^{m+1}$, so

$$\varrho(x_2, x_\infty) \leqslant c\varepsilon^{1+1/m}$$

which concludes the proof of Lemma 2.20 and so of Theorem 4.

Chapter 3. The estimation of certain kernels

We begin by studying the behavior of our families of balls under suitable mappings of the underlying space. The situation we study is motivated by the process of freeing of vector fields in Rothschild and Stein [RS].

Suppose as before that $\{Y_j\}$, $1\leq j\leq q$ is a family of C^∞ vector fields on a connected open set $\Omega\subset\mathbf{R}^N$. Assume that the vector field Y_j has formal degree d_j, that for any j and k we can write

$$[Y_j, Y_k] = \sum_{d_l\leq d_j+d_k} c_{jk}^l(x)\, Y_l, \quad c_{jk}^l \in C^\infty(\Omega) \tag{1}$$

and that the vector fields span the tangent space at each point of Ω.

Suppose in addition there is a connected open set $V\subset\mathbf{R}^M$ (with coordinates $s_1, \ldots, s_M$) and C^∞ vector fields $\tilde{Y}_1, \ldots, \tilde{Y}_q$ on $\Omega\times V$ with the following properties

(1) For each j, $1\leq j\leq q$,

$$\tilde{Y}_j(x, s) = Y_j(x) + \sum_{k=1}^{M} a_{jk}(x, s)\frac{\partial}{\partial s_k}.$$

(2) The formal degree of $\tilde{Y}_j$ is d_j and we can write for any j and k

$$[\tilde{Y}_j, \tilde{Y}_k] = \sum_{d_l\leq d_j+d_k} \tilde{c}_{jk}^l(x, s)\, \tilde{Y}_l$$

where $\tilde{c}_{jk}^l \in C^\infty(\Omega\times V)$.

(3) The vector fields $\tilde{Y}_1, \ldots, \tilde{Y}_q$ are linearly independent, and span the tangent space at each point of $\Omega\times V$.

The vector fields $Y_1, \ldots, Y_q$ give rise to a metric ϱ and a family of balls $\{B(x, \delta)\}$ on Ω. In the same way, the vector fields $\tilde{Y}_1, \ldots, \tilde{Y}_q$ give rise to a metric $\tilde{\varrho}$ and a family of balls $\{\tilde{B}((x, s), \delta)\}$ on $\Omega\times V$. We want to study the relationship between these metrics.

Let π and π_2 denote the projections from $\Omega\times V$ to Ω and V

$$\pi(x, s) = x$$

$$\pi_2(x, s) = s.$$

If $d\pi$ is the differential of this mapping, then condition (1) on the vector fields $\tilde{Y}_1, \ldots, \tilde{Y}_q$ is equivalent to

$$d\pi(\tilde{Y}_j) = Y_j, \quad 1\leq j\leq q.$$

This implies that if $\tilde{\varphi}: [0, 1] \to \Omega \times V$ is a curve satisfying

$$\tilde{\varphi}'(t) = \sum_{j=1}^{q} a_j(t)\, \tilde{Y}_j(\tilde{\varphi}(t))$$

and if we let $\varphi: [0, 1] \to \Omega$ be the composite

$$\varphi(t) = \pi(\tilde{\varphi}(t))$$

then

$$\varphi'(t) = \sum_{j=1}^{q} a_j(t)\, Y_j(\varphi(t)).$$

This observation leads immediately to the following.

LEMMA 3.1. (a) $\pi: \tilde{B}((x, s), \delta) \to B(x, \delta)$ *and this mapping is onto.*
(b) $\pi(\exp(\Sigma\, a_j \tilde{Y}_j)(x, s)) = \exp(\Sigma\, a_j Y_j)(x)$.
(c) *If* $x_1, x_2 \in \Omega$ *and* $s_1, s_2 \in V$ *then* $\varrho(x_1, x_2) \leq \tilde{\varrho}((x_1, s_1), (x_2, s_2))$.

Now fix $\Phi \in C_0^\infty(V)$. As in Rothschild-Stein [RS], we can define a restriction mapping from functions on $\Omega \times V$ to functions on Ω by the formula

$$RF(y) = \int_V F(y, s)\, \Phi(s)\, ds$$

for $y \in \Omega$, $F(y, s)$ a function on $\Omega \times V$. In order to estimate these restrictions for suitable F we need

LEMMA 3.2. *Let* $E \subset\subset \Omega$ *be a compact set. There is a constant* C *so that if* $x \in E$ *and if* $y \in B(x, \delta)$ *then*

$$\left| \int_V \chi_{\tilde{B}((x, 0), \delta)}(y, s)\, \Phi(s)\, ds \right| \leq C \frac{|\tilde{B}((x, 0), \delta)|}{|B(x, \delta)|}.$$

Proof. By Theorem 7, we can choose an N-tuple of vector fields $Y_{i_1}, \ldots, Y_{i_N}$ with the following properties: Let $U_j = Y_{i_j}$ for $1 \leq j \leq N$ and let $V_1, \ldots, V_{q-N}$ denote the remaining vector fields. For $u \in \mathbf{R}^N$, $v \in \mathbf{R}^{q-N}$, define

$$\Phi_v(u) = \exp(u \cdot U + v \cdot V)(x).$$

Then if $|v_j|<\eta_2\delta^{d(V_j)}$, $1\leq j\leq q-N$ the mapping

$$u\mapsto\Phi_v(u)$$

is globally one to one for $|u_j|<\eta_1\delta^{d(U_j)}$, and the image of this map contains the ball $B(x,\eta_2\delta)$.

We also consider the map $\tilde{\Phi}\colon\mathbf{R}^q\to\mathbf{R}^q$ given by

$$\tilde{\Phi}(u,v)=\exp(u\cdot U+v\cdot V)((x,0)).$$

Note that

$$\pi\circ\tilde{\Phi}(u,v)=\Phi_v(u).$$

Let $y\in B(x,0)$ and let

$$\Sigma_y=\{(u,v)\in\mathbf{R}^q|\,|u_j|<\eta_1\delta^{d(U_j)},\ |v_j|<\eta_2\delta^{d(V_j)},\text{ and }\Phi_v(u)=y\}.$$

Since the mapping Φ_v is globally one to one, for each v there is a unique $\theta(v)\in\mathbf{R}^N$ so that

$$(\theta(v),v)=\Sigma_y$$

i.e. $\Phi_v(\theta(v))=y$.

If we differentiate this last equation with respect to v_j and use the fact that the Jacobian of Φ_v is bounded between two positive multiples of $\lambda_I(x)$, it follows that

$$|J\theta(v)|\leq C|\lambda_I(x)|^{-1}.$$

Next consider the map $\chi\colon\mathbf{R}^{q-N}\to\mathbf{R}^{q-N}$ given by

$$\chi(v)=\pi_2\tilde{\Phi}(\theta(v),v).$$

Since $\tilde{\Phi}$ is a diffeomorphism

$$|J\chi|\leq C'|J\theta|\leq C''|\lambda_I(x)|^{-1}.$$

Now in the integral

$$\int_V\chi_{\tilde{B}((x,0),\eta\delta)}(y,s)\,\Phi(s)\,ds$$

we are just integrating Φ over the image of the map χ. The change of variables formula for multiple integrals now gives

$$\left|\int_V \chi\right| \leqslant C|\lambda_I(x)|^{-1}\delta^{\Sigma d(V_j)}| \leqslant C\frac{\delta^{\Sigma d(U_j)+d(V_j)}}{|\lambda_I(x)|\,\delta^{\Sigma d(U_j)}}$$

$$\leqslant C\frac{\tilde{B}((x,0),\delta)|}{|B(x,\delta)|}.$$

We turn now to the proof of Theorem 5.

Choose ε so small that Lemma 3.2 holds. We then write

$$k(x,0;y,s)=\sum_j k(x,0;y,s)\tilde{\chi}_j(x,0;y,s)$$

where

$$\tilde{\chi}_j(x,0;y,s)=\begin{cases}1 & \text{if } \varepsilon 2^{-j}\leqslant\tilde{\varrho}(x,0;y,s)<2\varepsilon\cdot 2^{-j}\\ 0 & \text{otherwise.}\end{cases}$$

Thus

$$|k(x,0;y,s)|\leqslant C\sum_{2^{-j}\geqslant\varrho(x,y)/2\varepsilon}\frac{(2\varepsilon 2^{-j})^{\alpha}}{\operatorname{Vol}\tilde{B}(x,0;\varepsilon 2^{-j})}\chi_{\tilde{B}(x,0,2\varepsilon\cdot 2^{-j})}(y,s)$$

where $\chi_{\tilde{B}}$ is the characteristic function of $\tilde{B}$. Now we apply Lemma 3.2 to obtain

$$\left|\int k(x,0;y,s)\,\varrho(s)\,ds\right|\leqslant C\sum_{\substack{j\\ 2^{-j}\geqslant\varrho(x,y)/2\varepsilon}}\frac{2^{-j\alpha}}{\operatorname{Vol}(B(x,2^{-j}))}.$$

We then obtain the estimate for Rk by comparing the sum to an integral. The estimation of $X_{i_1}\dots X_{i_j}Rk$ is similar. (See the discussion of $X_iRk-R\tilde{X}_jk$ in [RS] p. 302.)

Appendix: Exponential mappings and the Campbell-Hausdorff formula

In this appendix we briefly recall the basic properties of exponential mappings induced by vector fields. Thus suppose that $\Omega\subset\mathbf{R}^N$ is an open set, and that Y is a C^∞ real vector field defined on Ω. Y then induces a local one parameter group of transformations on Ω as follows: for each $x\in\Omega$ let $E_Y(t,x)=E(t,x)$ be the unique solution to the initial value problem

$$\frac{\partial E}{\partial t}(t,x)=Y(E(t,x)),\quad E(0,x)=x \tag{1}$$

This unique solution always exists for $|t|$ sufficiently small. In fact, if $K \subset\subset \Omega$ is compact, there is a number $\delta_0 > 0$ so that $E: (-\delta_0, \delta_0) \times K \to \Omega$ is a C^∞ mapping. Moreover, if the vector field $Y = Y(u_1, \ldots, u_p)$ depends in a C^∞ manner on a parameter $u = (u_1, \ldots, u_p) \in U \subset \mathbf{R}^p$ where U is open, then if $K \subset\subset \Omega$ and $L \subset\subset U$ are compact, there is a constant $\delta_0 > 0$ so that the solution

$$E(u_1, \ldots, u_p, t, x) = E\colon L \times (-\delta_0, \delta_0) \times K \to \Omega$$

is a C^∞ mapping. (A discussion of existence and uniqueness of solutions of (1), and of smoothness of dependence on parameters can be found in Dieudonné [D], Chapter X.)

The uniqueness of the solution to (1) shows that

$$E(s, E(t, x)) = E(s+t, x) \quad \text{if } x \in K,\ |s|+|t| < \delta_0 \tag{2}$$

and

$$E_Y(\lambda t, x) = E_{\lambda Y}(t, x) \quad \text{if } x \in K,\ |\lambda t| < \delta_0. \tag{3}$$

Equation (2) shows that the mapping

$$x \mapsto E(t, x)$$

is a diffeomorphism on compact subsets of Ω if t is sufficiently small. Equation (3) shows that $E_{ty}(1, x)$ is well defined on compact subsets for all t sufficiently small. We define

$$\exp(Y)(x) = E_y(1, x)$$

whenever the right hand side is defined. In particular, $\exp(tY)(x)$ is always defined for small $|t|$.

If f is a C^∞ function defined near a point $x_0 \in \Omega$ the function

$$F(t) = f(\exp(tY)(x_0))$$

is defined and C^∞ near $t=0$, and the differential equation (1) shows that

$$F'(t) = (Yf)(\exp(tY)(x_0)).$$

This shows that the Taylor series expansion of F at $t=0$ is given by

$$F(t) \sim \sum_{n=0}^{\infty} \frac{1}{n!} (Y^n f)(x_0)\, t^n = e^{tY} f(x_0), \tag{4}$$

where the last expression is thought of as a formal power series in (tY).

Now suppose that $Y_1, \ldots, Y_P$ are real C^∞ vector fields on Ω. If $u=(u_1, \ldots, u_P) \in \mathbf{R}^P$, then $\Sigma_{j=1}^P u_j Y_j$ is again a real C^∞ vector field on Ω, and we may consider $\exp(\Sigma_{j=1}^P u_j Y_j)(x)$ for $x \in \Omega$, $u \in \mathbf{R}^P$. It is clear that this is well defined for

$$|u| = \left(\sum_{j=1}^{P} |u_j|^2 \right)^{1/2}$$

sufficiently small. In fact if $K \subset\subset \Omega$ is compact, there exists $\delta_0 > 0$ so that $\exp(\Sigma u_j Y_j)(x)$ exists for $x \in K$ and $|u| < \delta_0$. Moreover the mapping

$$(u_1, \ldots, u_P; x) \mapsto \exp\left(\sum u_j Y_j \right)(x)$$

is infinitely differentiable. For $x \in \Omega$, we may define

$$\exp_x(u_1, \ldots, u_P) = \exp\left(\sum u_j Y_j \right)(x).$$

Then $\exp_x$ is a C^∞ mapping of a neighborhood of 0 in $\mathbf{R}^P$ to Ω with $\exp_x(0)=x$. This induces the differential mapping $d(\exp_x)$: $T_0\mathbf{R}^P \to T_x\Omega$ between tangent spaces. It follows easily from the differential equation (1) that

$$d(\exp_x)\left(\frac{\partial}{\partial u_j} \right) = Y_j(x), \quad j = 1, \ldots, P.$$

In particular, if $P=N$ and $Y_1, \ldots, Y_N$ is a basis for the tangent space $T_x\Omega$, this shows that $d(\exp_x)$ is a bijection. The inverse function theorem, together with the smoothness of the mapping $(u_1, \ldots, u_N; x) \mapsto \exp(\Sigma u_j Y_j)(x)$, proves:

PROPOSITION 4.1. *Let $Y_1, \ldots, Y_N$ be C^∞ real vector fields on an open set $\Omega \subset \mathbf{R}^N$. Suppose $Y_1, \ldots, Y_N$ are linearly independent at each point of Ω. let $K \subset\subset \Omega$ be compact. Then there are finite positive constants δ_1, δ_2, and C so that:*

(a) *For $x \in K$, $\exp_x$ is a diffeomorphism of $\{u \in \mathbf{R}^N \mid |u| < \delta_1\}$ onto a neighborhood V_x of x in Ω.*

(b) *For $x \in K$, $\{y \in \mathbf{R}^N \mid |x-y| < \delta_2\} \subset V_x$.*

(c) *For $x \in K$, if $u_j \in \mathbf{R}^N$ with $|u_j| < \delta_1$ for $j=1,2$,*

$$|\exp_x(u_1) - \exp_x(u_2)| \leq C|u_1 - u_2|.$$

COROLLARY 4.1. *Let $x_0 \in \Omega$. Then there is an open neighborhood U of x_0 and constants $\delta_0 > 0$, $C < \infty$ so that*

(a) *If* $x\in U$, $\exp_x$ *is a diffeomorphism of* $\{u\in\mathbf{R}^N \mid |u|<\delta_0\}$ *onto an open set in* Ω *containing* U.

(b) *If* $x, y\in U$ *and* $y=\exp_x(u)$ *with* $|u|<\delta_0$ *then* $\exp_x(tu)\in\Omega$ *for* $0\leq t\leq 1$.

(c) *If* $x\in U$, $u_j\in\mathbf{R}^N$ *with* $|u_j|<\delta_0$ *for* $j=1,2$ *then*

$$|\exp_x u_1-\exp_x u_2|\leq c|u_1-u_2|.$$

We shall need a generalization of (a) which is proved in the same way:

COROLLARY 4.2. *Suppose* $V_1, \ldots, V_r$ *are some additional vector fields. Then if* $v=(v_1, \ldots, v_r)$ *with* $|v_j|<\delta_0$

$$\exp_x(u_1 Y_1+\ldots+u_N Y_N+v_1 V_1+\ldots+v_r V_r)$$

is a diffeomorphism of $|u|<\delta_0$ *onto an open subset of a neighborhood of*

$$\exp_x(v_1 V_1+\ldots+v_r V_r).$$

We shall call U a *normal neighborhood* of x_0. If U is a normal neighborhood of x_0 then for each $x\in U$ we can introduce coordinates in U via the mapping $(\exp_x)^{-1}$: $U\to\mathbf{R}^N$. We call these the *canonical coordinates* at x, relative to the linearly independent vector fields $Y_1, \ldots, Y_N$.

Next, we return to the more general situation of P (not necessarily linearly independent) vector fields $Y_1, \ldots, Y_P$ on Ω. If $x_0\in\Omega$ and f is a C^∞ function defined near x_0 the function

$$F(u_1, \ldots, u_P)=f\left(\exp\left(\sum u_j Y_j\right)(x_0)\right)$$

is defined and C^∞ near $0\in\mathbf{R}^N$. As a generalization of equation (4) we have

PROPOSITION 4.2. *The formal Taylor series of* F *at* $0\in\mathbf{R}^P$ *is given by*

$$F(u_1, \ldots, u_P)\sim\sum_{n=0}^{\infty}\frac{1}{n!}\left(\sum u_j Y_j\right)^n f(x_0)\equiv e^{\Sigma u_j Y_j}f(x_0).$$

Proof. Define

$$G(t, u_1, \ldots, u_P)=F(tu_1, \ldots, tu_P).$$

Then $G(t, u_1, \ldots, u_P)=f(\exp t(\Sigma u_j Y_j)(x_0))$, and from equation (4) we have

$$\frac{\partial^n G}{\partial t^n}(0, u_1, \ldots, u_P)=\left(\sum u_j Y_j\right)^n f(x_0).$$

On the other hand

$$\frac{\partial^n G}{\partial t^n}(0, u_1, \ldots, u_P) = \sum_{|\alpha|=n} \frac{n!}{\alpha!} u^\alpha \frac{\partial^\alpha F}{\partial u^\alpha}(0)$$

so

$$\sum_{|\alpha|=n} \frac{u^\alpha}{\alpha!} \frac{\partial^\alpha F}{\partial u^\alpha}(0) = \frac{1}{n!}\left(\sum u_j Y_j\right)^n f(x_0). \qquad \text{Q.E.D.}$$

We have seen that if $K \subset\subset \Omega$ is compact, there exists $\delta_0 > 0$ so that the mapping $K \to \Omega$ given by

$$x \mapsto \exp\left(\sum u_j Y_j\right)(x)$$

is a diffeomorphism if $|u| < \delta_0$. We shall often have need to compose two such diffeomorphisms. Let $x_0 \in \Omega$ and let f be a C^∞ function defined near x_0. Then the function

$$F(s_1, \ldots, s_P, t_1, \ldots, t_P) = f\left(\exp\left(\sum s_j Y_j\right) \circ \exp\left(\sum t_j Y_j\right)(x_0)\right)$$

is defined and C^∞ near $0 \in \mathbf{R}^{2P}$. According to Proposition 2, for $|t|$ sufficiently small

$$\sum_{|\alpha|=n} \frac{s^\alpha}{\alpha!} \frac{\partial^{|\alpha|} F}{\partial s^\alpha}(0, t) = \frac{1}{n!}\left(\left(\sum s_j Y_j\right)^n f\right)\left(\exp\left(\sum t_j Y_j\right)(x_0)\right).$$

Applying the proposition again we have

$$\sum_{\substack{|\alpha|=n \\ |\beta|=m}} \frac{s^\alpha t^\beta}{\alpha!\beta!} \frac{\partial^{|\alpha|+|\beta|} F}{\partial s^\alpha \partial t^\beta}(0,0) = \frac{1}{m!}\frac{1}{n!}\left(\sum t_j Y_j\right)^m \left(\sum s_j Y_j\right)^n f(x_0).$$

Thus the formal power series of F at $(0,0)$ is given by

$$F(s_1, \ldots, s_P, t_1, \ldots, t_P) \sim \sum_{m,n=0}^{\infty} \frac{1}{m!\,n!}\left(\sum t_j Y_j\right)^m \left(\sum s_j Y_j\right)^n f(x_0).$$

But the right hand side, viewed as a formal series in t and s is exactly $e^{\Sigma t_j Y_j} e^{\Sigma s_j Y_j} f(x_0)$. According to the Campbell-Hausdorff formula, the formal series $e^{\Sigma t_j Y_j} e^{\Sigma s_j Y_j}$ can be written as e^Z where

$$Z=\sum_{n=1}^{\infty}(-1)^{n+1}n^{-1}\sum_{\alpha_i+\beta_i\neq 0}(\alpha!\beta!(\alpha+\beta))^{-1}\left(\mathrm{ad}\left(\sum t_j Y_j\right)\right)^{\alpha_1}\left(\mathrm{ad}\left(\sum s_j Y_j\right)\right)^{\beta_1}$$
$$\times\ldots\left(\mathrm{ad}\left(\sum s_j Y_j\right)^{\beta_n-1}S\right) \tag{5}$$

(see Hörmander [H], p. 160, or Hochschild [Ho], Chapter X). If M is a positive integer, let Z_M be the *finite* partial sum of the formal series for Z so that $Z-Z_M=O(|s|^M+|t|^M)$. We now have

$$f\left(\exp\left(\sum s_j Y_j\right)\circ\exp\left(\sum t_j Y_j\right)(x_0)\right)-f(\exp(Z_M)(x_0))=O(|s|^M+|t|^M).$$

Since f is arbitary, we also have:

PROPOSITION 4.3. *Let $K\subset\subset\Omega$ be compact, and let M be a positive integer. There exists a constant C so that for $x\in K$*

$$\left|\exp\left(\sum s_j Y_j\right)\circ\exp\left(\sum t_j Y_j\right)(x)-\exp Z_m(x)\right|\leq C(|s|^M+|t|^M).$$

References

[C] CARATHÉODORY, C., Untersuchungen über die Grundlagen der Thermodynamik. *Math. Ann.*, 67 (1909), 355–386.

[Ch] CHOW, W. L., Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung. *Math. Ann.*, 117 (1940), 98–115.

[D] DIEUDONNÉ, J., *Foundations of modern analysis*. Academic Press, New York, 1960.

[FP] FEFFERMAN, C. & PHONG, D. H., Subelliptic eigenvalue problems, in *Conference on Harmonic Analysis in honor of Antoni Zygmund*, vol. 2, pp. 590–606. Wadsworth, 1983.

[FH] FOLLAND, G. & HUNG, H. T., Non-isotropic Lipschitz spaces, in *Harmonic Analysis in Euclidean Spaces*. Amer. Math. Soc., Part 2, pp. 391–394. Providence, 1979.

[Ho] HOCHSCHILD, G., *The structure of Lie groups*. Holden-Day Inc., San Francisco, London, Amsterdam, 1965.

[H] HÖRMANDER, L., Hypoelliptic second order differential equations. *Acta Math.*, 119 (1967), 147–171.

[NS] NAGEL, A. & STEIN, E. M., *Lectures on pseudo-differential operators*. Math. Notes Series no. 24. Princeton Univ. Press, Princeton, 1979.

[NSW] NAGEL, A., STEIN, E. M. & WAINGER, S., Boundary behavior of functions holomorphic in domains of finite type. *Proc. Nat. Acad. Sci. U.S.A.*, 78 (1981), 6596–6599.

[RS] ROTHSCHILD, L. P. & STEIN, E. M., Hypoelliptic differential operators and nilpotent groups. *Acta Math.*, 137 (1976), 247–320.

[Sa] SANCHEZ, A., Estimates for kernels associated to some subelliptic operators. Thesis, Princeton University, 1983.

[St1] STEIN, E. M., *Singular integrals and differentiability properties of functions*. Princeton Univ. Press, Princeton, 1970.

[St2] — *Boundary behavior of holomorphic functions of several complex variables*. Math. Notes Series no. 11. Princeton Univ. Press, Princeton, 1972.

Received November 11, 1983