HYPOELLIPTIC DIFFERENTIAL OPERATORS AND NILPOTENT GROUPS

BY

LINDA PREISS ROTHSCHILD[1] and E. M. STEIN[1]

Institute for Advanced Study
University of Wisconsin
Madison, Wisc., USA

Princeton University
Princeton, N. J., USA

Contents

[1] Supported in part by National Science Foundation grants MPS72-05055 A02 and GP36318, respectively.

§ 1. Introduction

We shall deal with a general class of second-order hypoelliptic partial differential equations. For this purpose we develop an appropriate class of singular integral operators; these will be modeled on convolution operators on certain nilpotent Lie groups. As a result we are able to construct parametrices and to obtain sharp regularity results in terms of various Sobolev spaces and Lipschitz spaces.

Background. The first example of the kind of differential operators to be studied here came from several complex variables. The operator is the $\Box_b$ Laplacian associated with the boundary analogue of the $\bar\partial$ complex. Its highest order part is a quadratic expression in certain vector fields, which are the real and imaginary parts of the tangential holomorphic vector fields. These vector fields do not span the tangent space. Kohn [13] was able to show, nevertheless, that under appropriate geometric hypotheses $\Box_b$ is hypoelliptic because the missing directions arise as commutators of the complex tangential directions.

A far-reaching generalization of this basic idea was obtained by Hörmander [10]. He considered the operator

$$X_0+\sum_{j=1}^{n} X_j^2, \tag{1.1}$$

$X_0, \ldots, X_n$ are $n+1$ smooth real vector fields on a manifold M with the property that the commutators up to a certain order suffice to span the tangent space at each point. Hörmander proved that the operator (1.1) is then hypoelliptic. Alternative treatments of (1.1) were later obtained by Kohn [15] and Radkevitch [20]. All the arguments given were essentially L^2 in character, but even in this context did not usually give optimal estimates. The problem arose, therefore, of constructing an appropriate class of operators which could be used to find approximate inverses to (1.1) and in terms of which sharp estimates in L^p spaces, Lipschitz spaces, etc. could be made.

The first step in this direction was taken in Folland and Stein [8], where the $\Box_b$ Laplacian was studied in terms of operators modeled on convolution operators on the Heisenberg group. Parametrices were constructed and various sharp estimates were obtained for $\Box_b$. The intervention of the Heisenberg group (the simplest nilpotent, non-abelian Lie group) in this context should not have been surprising. It suffices to recall that this group is naturally isomorphic with the boundary of the unit ball in $\mathbf{C}^n$, $n>1$, in its unbounded realization.

Outline of the paper. The main objectives of the present paper are (i) to find nilpotent groups suitable to the analysis of the operators (1.1) or variants thereof; (ii) to study related questions concerning operators on these groups; (iii) to apply this analysis to the original partial differential operators in order to obtain the regularity results described below.

We sketch first the idea leading to the construction of the suitable nilpotent groups. Let $X_1, ..., X_n$ be given vector fields on M with the property that the commutators up to length not exceeding r span the tangent space at each point, and consider the following examples:

(a) The X_j are linearly independent and span the tangent space (i.e. $r=1$). The nilpotent group is then $\mathbf{R}^n$, and the singular integrals are the standard ones.

(b) $M=\mathbf{R}^2=\{(x, y)\}$, $X_1=\partial/\partial x$ and $X_2=x(\partial/\partial y)$. Observe that $[X_1, X_2]=\partial/\partial y$ so that X_1 and $[X_1, X_2]$ span. Unfortunately, there exist no two-dimensional non-abelian nilpotent Lie groups. However, if we add an extra variable, t, and write $\tilde{X}_1=\partial/\partial x$, $\tilde{X}_2=\partial/\partial t+x(\partial/\partial y)$ then $\tilde{X}_1$, $\tilde{X}_2$ and $[\tilde{X}_1, \tilde{X}_2]$ span the Lie algebra of the three-dimensional Heisenberg. Observe also that once one proves that $\tilde{X}_1^2+\tilde{X}_2^2$ is hypoelliptic, then the hypoellipticity of $X_1^2+X_2^2$ follows as an easy consequence. It is clear from this example that the dimension of the group to be used may be higher than that of the given manifold M.

(c) Take M to be the boundary of a smooth domain in $\mathbf{C}^{l+1}$, $n=2l$, and the vector fields are the real and imaginary parts of holomorphic vector fields which are tangent to M. If the Levi form is everywhere nondegenerate, the Heisenberg group is the appropriate one; this is the case studied in [8]. However, if the signature of the Levi form varies, the group that one might naively associate would vary from point to point. In order to deal with this situation, as well as the more general situation where this phenomenon occurs in more complicated forms, we lift to a larger group for which all possible varying structures occur as quotients. This group is the free nilpotent group of step r. (In our specific example above, $r=2$.)

We now leave these examples and return to our general problem. We shall show that by adding an appropriate number of new variables we can lift our original vector fields to an extended space $\tilde{M}$. The resulting vector fields $\tilde{X}_1, ..., \tilde{X}_n$, and their commutators up to length r are now free, i.e. satisfy the minimal number of relations at a given point, and span the tangent space of $\tilde{M}$. At the same time we prove that at each point the vector fields $\tilde{X}_1, ..., \tilde{X}_n$ are well approximated by the left-invariant vector fields $Y_1, ..., Y_n$ which generate the Lie algebra of the free nilpotent group N_F of step r.

Thus the study of the operator, say

$$\sum_{j=1}^{n} X_j^2 \tag{1.2}$$

is reduced to the analysis of the operator $\sum_{j=1}^n \tilde{X}_j^2$, and the latter is linked to the left-invariant operator $\sum_{j=1}^n Y_j^2$ on a (free) nilpotent Lie group.

The results concerning the addition of variables to vector fields and the approximation by left-invariant vector fields on free groups are obtained as Theorems 4 and 5 respectively in Part II. The proofs of these theorems are intertwined and depend on a complicated inductive procedure. We shall not describe the details here, but suggest that the reader look at the statements of Lemmas 8.2, 8.3, 8.4, and 8.5 for the main ideas of the induction.

In the course of the proof of these theorems we exhibit a basic mapping Θ of $\tilde{M} \times \tilde{M}$ to N_F. Among other things it generalizes the function $(\xi, \eta) \to \xi - \eta$. (Indeed, in example (a) Θ is essentially given by this function.)

In addition to studying operators of the form (1.1) we shall also consider the operators

$$\mathcal{L} = \sum_{j=1}^{n} X_j^2 + \tfrac{1}{2} \sum_{j,k} c_{jk}[X_j, X_k] \tag{1.3}$$

where the c_{jk} are given skew symmetric matrices of smooth functions. We are led to deal with this class of operators because the $\Box_b$ Laplacian can be written as (1.3), modulo lower order terms, where the c_{jk} are in fact matrices. (Recall that $\Box_b$ acts on (p, q)-forms.)

Part I of this paper is devoted to the study of operators of the form (1.3) on groups, where the X_j are left-invariant vector fields and the c_{jk} are assumed to be constant. In the case where c_{jk} are actually scalars we give sufficient conditions for the hypoellipticity of $\mathcal{L}$ in terms of the size of the imaginary part of the c_{jk} (Theorems 1 and 1′). We also show, using the rudiments of Fourier analysis on nilpotent groups, that the conditions are necessary for a large class of groups. (See Theorem 2.) When the nilpotent groups have a natural homogeneous structure (e.g. are graded) then whenever $\mathcal{L}$ is hypoelliptic it has a unique fundamental solution which is homogeneous in the appropriate sense. It turns out (Theorem 3) that this fundamental solution then depends smoothly on the parameters c_{jk}.(1)

Part III is devoted to the analysis of the analogue of (1.3) in terms of the free vector fields. We consider

$$\tilde{\mathcal{L}} = \sum_{j=1}^{n} \tilde{X}_j^2 + \tfrac{1}{2} \sum_{j,k} c_{jk}[\tilde{X}_j, \tilde{X}_k]. \tag{1.4}$$

We construct a parametrix for $\tilde{\mathcal{L}}$ as follows. For each $\xi \in \tilde{M}$, we let $k_\xi(\cdot)$ be the fundamental solution of the operator $\mathcal{L}_\xi = \sum_{j=1}^{n} Y_j^2 + \frac{1}{2}\sum_{j,k} c_{jk}(\xi)[Y_j, Y_k]$ on the free nilpotent group N_F. Then the kernel of a parametrix for (1.4) is (with small modifications) given by $K(\xi, \eta) = k_\xi(\Theta(\eta, \xi))$.

The parametrix and the resulting regularity properties can then be studied by following

(1) At this stage our work depends on some results of Folland [6].

the techniques previously used in [8]. It might be worthwhile, however, to call to the reader's attention some of the features of this analysis which are not straightforward adaptations from [8]. First, the properties related to the pseudo-metric which is determined by Θ (see § 12) are more complicated and a require a different approach to prove. Next, the question of differential operators acting on operators such as the parametrix is dealt with by a more direct, computational method. (See § 14.) Finally, the problem of bounds for classical Lipschitz spaces (Theorem 14, § 16) had not been considered in [8].(1)

Part IV deals with the main applications. The passage from $\tilde{M}$ back to M (and from (1.4) to (1.3)) is accomplished by a simple technique which amounts to integration in the added variables. This leads to the main regularity theorem for solutions of (1.3). An example of this result (Theorem 16 in § 17) is as follows. Suppose $\mathcal{L}(f)=g$, and g belongs to the fractional Sobolev space $L^p_\alpha(M)$, $1<p<\infty$, $\alpha\geqslant 0$. Then locally f belongs to $L^p_{\alpha+(2/r)}(M)$. (Recall that r is the least length required to obtain spanning commutators.) There are also further results of this kind for new Sobolev spaces which take into account the special directions $X_1, \ldots, X_n$, and also for the standard Lipschitz spaces.

In § 18 we show how the analysis must be modified to apply to the operators of type (1.1). The main difference arises in the choice of an appropriate nilpotent group to be used. The group we shall use has generators $Y_0, Y_1, \ldots, Y_n$; however, in determining the notion of "length" of a commutator, Y_0 is given twice the weight that is given to $Y_1, Y_2, \ldots, Y_n$.

Finally in § 19 we obtain the desired results for the $\square_b$ Laplacian. Thus the estimates of [8] are generalized to the setting where the Levi form need not be non-degenerate and the metric is not necessarily one of the special metrics used in [8].

It is a pleasure to thank G. B. Folland, C. D. Hill, and J. J. Kohn for several useful discussions during the preparation of this paper.

Part I. Operators on nilpotent Lie groups

§ 2. Sufficient conditions for hypoellipticity

Let $\mathfrak{G}$ be a (finite-dimensional) real Lie algebra, and G a corresponding connected Lie group with exponential map $\exp \mathfrak{G}\to G$. Every $Y\in\mathfrak{G}$ acts as a left-invariant vector field (hence differential operator) on $\mathfrak{G}$ by the equation

$$(Yf)(x)=d/dt(f(x\exp tY))\big|_{t=0},$$

for $x\in G$ and $f\in C_0^\infty(G)$. (C_0^∞ denotes the space of smooth, compactly supported functions.) Any polynomial $\mathcal{L}$ on $\mathfrak{G}$ is therefore also a differential operator. We shall be concerned

(1) In this connection see Greiner and Stein [9].

with the problem of giving sufficient conditions on such a second order operator $\mathcal{L}$ that it be hypoelliptic in the following sense. Suppose U is an open subset of G and u is a distribution on U so that $\mathcal{L}(u) \in C^\infty(U)$; then it follows that $u \in C^\infty(U)$.

For $\varphi, \psi \in C_0^\infty(G)$ we let $(\varphi, \psi) = \int_G \varphi\bar{\psi}dx$, where dx is right-invariant Haar measure(1) on G, and set $\|\varphi\| = (\varphi, \varphi)^{\frac{1}{2}}$. We observe that $(Y\varphi, \psi) = -(\varphi, Y\psi)$, for any $Y \in \mathfrak{G}$ whenever $\varphi, \psi \in C_0^\infty(G)$. If $c = (c_{jk})$ is any $n \times n$ matrix, we denote by $\|c\|$ the usual norm of the operator acting on the standard n-dimensional coordinate Hilbert space.

THEOREM 1. *Suppose $Y_1, Y_2, \ldots, Y_n$ generate the Lie algebra $\mathfrak{G}$. Let $\mathcal{L}$ be the left-invariant differential operator on G given by*

$$\mathcal{L} = \sum_{j=1}^{n} Y_j^2 + \frac{i}{2} \sum_{j,k} b_{jk}[Y_j, Y_k],$$

where $(b_{jk}) = b$ is a real skew-symmetric matrix. Then $\mathcal{L}$ is hypoelliptic if

$$\|b\| < 1.$$

This theorem is proved by simple adaptations of arguments of Hörmander [10], Kohn [14], [15], and Radkevich [20]. The argument can be modified so as to cover the general case (i.e. vector fields on some manifold whose commutators up to a certain length span the tangent space, and not necessarily those that arise from a finite-dimensional Lie algebra). However, this general case will anyway be subsumed in our later considerations where we will make more precise statements.

The proof of Theorem 1 may be generalized to take into account the situation that arises (i) when the matrix b is complex, (ii) when there are linear relations among the commutators $[Y_j, Y_k]$. To express the result we need the following additional definitions. Suppose ϱ is any real skew-symmetric matrix. Then under conjugation by an orthogonal matrix, ϱ can be put in block diagonal form. Each block is of the form $\begin{pmatrix} 0 & \varrho_j \\ -\varrho_j & 0 \end{pmatrix}$, where ϱ_j is real. There are $n/2$ such blocks if n is even and $(n-1)/2$ if n is odd. In the latter case a zero appears in the last diagonal entry. One should observe that the numbers $\pm i\varrho_j$, $j = 1, 2, \ldots, [n/2]$, are the eigenvalues of ϱ. We define the "trace norm," $\|\varrho\|_1$, of ϱ as the sum of the absolute values of the eigen-values of ϱ, i.e. $\|\varrho\|_1 = 2\sum_{j=1}^{[n/2]} |\varrho_j|$.

The following facts are obvious as soon as ϱ is put into block-diagonal form. Here b and ϱ range over real $n \times n$ skew-symmetric matrices.

---

(1) The distinction between right- and left-invariant Haar measure will be irrelevant in most of what follows since all nilpotent groups are unimodular.

$$|\operatorname{tr}(b\varrho)| \leq \|b\|\,\|\varrho\|_1, \tag{2.1}$$

$$\|b\| = \sup_{\|\varrho\|_1 \leq 1} |\operatorname{tr}(b\varrho)|; \quad \|\varrho\|_1 = \sup_{\|b\| \leq 1} |\operatorname{tr}(b\varrho)|. \tag{2.2}$$

The generalization of Theorem 1 is as follows. We consider $\mathcal{L} = \sum_{j=1}^n Y_j^2 + \frac{1}{2}\sum_{j,k} c_{jk}[Y_j, Y_k]$ where $(c_{jk}) = c = a + ib$ and a and b are real $n \times n$ skew-symmetric matrices. Now let S be the subspace of all real skew-symmetric matrices $s = (s_{jk})$ so that $\sum_{j,k} s_{jk}[Y_j, Y_k] = 0$. We let R be the subspace of real skew-symmetric matrices spanned by a and S. Finally, let $R^\perp$ denote the orthogonal complement of R with respect to the inner product $(d_1, d_2) = -\operatorname{tr}(d_1 d_2)$, on the real skew-symmetric matrices.

THEOREM 1'. *Let $\mathcal{L} = \sum Y_j^2 + \frac{1}{2}\sum_{j,k} c_{jk}[Y_j, Y_k]$. Then $\mathcal{L}$ is hypoelliptic if*

$$\sup_{\substack{\|\varrho\|_1 \leq 1 \\ \varrho \in R^\perp}} |\operatorname{tr}(b\varrho)| < 1. \tag{2.3}$$

The proof of Theorem 1' (and hence of Theorem 1) requires two simple lemmas.

(2.4) LEMMA. *Suppose $\varrho_{jk} = i([Y_j, Y_k]f, f)$, for some $f \in C_0^\infty(G)$. The matrix ϱ is real and skew symmetric. Moreover,*

$$\|\varrho\|_1 \leq 2 \sum_{j=1}^n \|Y_j f\|^2. \tag{2.5}$$

Proof. The fact that ϱ_{jk} is real follows immediately from the fact that $(Zf, f) = -(f, Zf)$, for any real left-invariant vector field Z. Skew-symmetry is obvious. Now let $d = (d_{jk})$ be the matrix of any orthogonal transformation, and set $Y'_j = \sum_k d_{jk} Y_k$. Then clearly $\sum_j \|Y'_j f\|^2 = \sum_j \|Y_j f\|^2$. Moreover, if $\varrho'_{jk} = i([Y'_j, Y'_k]f, f)$, a simple computation shows that $\varrho' = d\varrho d^{-1}$. Thus with an appropriate choice of d we can assume that ϱ has been reduced to block diagonal form. It then suffices to prove the inequality corresponding to (2.5) for each block and then add these inequalities. Thus we need to show that:

$$|([Y_1, Y_2]f, f)| \leq \|Y_1 f\|^2 + \|Y_2 f\|^2. \tag{2.6}$$

So consider the fact that

$$((Y_1 + iY_2)f, (Y_1 + iY_2)f) \geq 0.$$

This means that $\|Y_1 f\|^2 + \|Y_2 f\|^2 + i\{(Y_2 f, Y_1 f) - (Y_1 f, Y_2 f)\} \geq 0$. However

$$((Y_2 f, Y_1 f) = -(Y_1 Y_2 f \;\; f) \quad \text{and} \quad (Y_1 f, Y_2 f) = -(Y_2 Y_1 f, f).$$

Hence $i([Y_1, Y_2]f, f) \leq \|Y_1 f\|^2 + \|Y_2 f\|^2$. A similar result holds with Y_1 and Y_2 interchanged giving (2.6) and proving the lemma.

(2.7) LEMMA. *Write $a=a^{(1)}+a^{(2)}$ with $a^{(1)}\in S^{\perp}\cap R$ and $a^{(2)}\in S\cap R$. If b satisfies* $\sup_{\|\varrho\|_1\leq 1,\, \varrho\in R^{\perp}} |\operatorname{tr}(b\varrho)|\leq\theta$, *then there exists $\gamma\in\mathbf{R}$ such that $b'=b+\gamma a^{(1)}$ satisfies*

$$\sup_{\substack{\|\varrho\|\leq 1\\ \varrho\in S^{\perp}}} |\operatorname{tr}(b'\varrho)|\leq\theta. \tag{2.8}$$

Proof. Identify b with the linear functional b^* on the vector space $R^{\perp}$, given by $b^*(\varrho)=-\operatorname{tr}(b\varrho)$, $\varrho\in R^{\perp}$. Then by the hypothesis,

$$|b^*(\varrho)|\leq\theta\|\varrho\|_1$$

for all $\varrho_1\in R^{\perp}$. Since $\|\varrho\|_1+\|\varrho+\varrho'\|_1\leq\|\varrho'\|_1$ (see (2.2) for any real skew-symmetric matrices ϱ,ϱ', the Hahn-Banach theorem guarantees the existence of a linear functional b'^* on $S^{\perp}\supset R^{\perp}$ such that

$$|b'^*(\varrho)|\leq\theta\|\varrho\|_1,\quad \text{all } \varrho\in S^{\perp}, \tag{2.9}$$

and

$$b'^*(\varrho)=b^*(\varrho)\quad \text{for } \varrho\in R^{\perp}. \tag{2.10}$$

Since $S^{\perp}$ is spanned by $R^{\perp}$ and a^1, $a^1\in S^{\perp}$, (2.10) implies that b'^* may be identified with a skew-symmetric matrix $b'=b+\gamma a^1$ for some real γ. Then (2.8) is immediate from (2.9), proving the lemma.

To prove the theorem we show first that we have the inequality

$$\sum_{j=1}^{n}\|Y_j f\|^2\leq A_c|(\mathcal{L}f,f)|,\ f\in C_0^{\infty}(G). \tag{2.11}$$

Write

$$(\mathcal{L}f,f)=((\sum Y_j^2)f,f)+\tfrac{1}{2}\sum_{j,k} b'_{jk}([Y_j,Y_k]f,f) \tag{2.12}$$
$$+\left(\frac{1-i\gamma}{2}\right)\sum_{j,k}a_{jk}([Y_j,Y_k]f,f)+\frac{i\gamma}{2}\sum_{j,k}a_{jk}^{(2)}([Y_j,Y_k]f,f).$$

This formula follows from the definition of $\mathcal{L}$ since $b'=b+\gamma a^{(1)}=b+\gamma(a-a^{(2)})$. Using $(Y_j^2 f,f)=-\|Y_j f\|^2$ we obtain from (2.12) the inequality

$$\sum\|Y_j f\|^2\leq|(\mathcal{L}f,f)|+\tfrac{1}{2}|\operatorname{tr}(b'\varrho)|+\frac{|1-i\gamma|}{2}|\operatorname{tr}(a\varrho)|+\frac{|\gamma|}{2}|\operatorname{tr}(a^{(2)}\varrho)|. \tag{2.13}$$

Note first that since $i([Y_j,Y_k]f,f)$ is real, $\operatorname{Im}(\mathcal{L}f,f)=-\tfrac{1}{2}\operatorname{tr}(a\varrho)$ and therefore

$$|\operatorname{tr}(a\varrho)|\leq 2|(\mathcal{L}f,f)|.$$

Furthermore $\sum \varrho_{jk} s_{jk} = 0$ whenever $\sum s_{jk}[Y_j, Y_k] = 0$. Thus $\varrho \in S^\perp$, and hence

$$\tfrac{1}{2}|\operatorname{tr}(b'\varrho)| \leqslant \theta \|\varrho\|_1/2 \leqslant \theta \sum \|Y_j f\|^2,$$

where $\theta = \sup_{\varrho\in S^\perp,\ \|\varrho\|_1\leqslant 1} |\operatorname{tr}(b'\varrho)| < 1$, by Lemmas 2.4 and 2.7. Also, since $a^{(2)} \in S$, $\operatorname{tr}(a^{(2)}\varrho) = 0$. Substituting these into (2.13) we obtain

$$\sum \|Y_j f\|^2 \leqslant |(\mathcal{L}f, f)| + \theta \sum \|Y_j f\|^2 + |1 - i\gamma|\,|(\mathcal{L}f, f)|.$$

This proves (2.11) with $A_c = (1 + |1 - i\gamma|)/(1-\theta)$.

Once (2.11) is proved one can follow the known arguments to prove hypoellipticity. (See e.g. Kohn [15].) Using Sobolev norms $\|\ \|_s = \|\ \|_{L^2_s}$ for functions supported in any fixed coordinate patch U of G, one first shows that (2.11) implies

$$\|\varphi\|^2_{L^2_\varepsilon} \leqslant C(|(\mathcal{L}\varphi, \varphi)| + \|\varphi\|^2),\quad \varphi \in C_0^\infty(U), \tag{2.14}$$

for some $\varepsilon > 0$. Next one shows that (2.14) implies the local estimate

$$\|\zeta\theta\|_{s+\varepsilon} \leqslant C\{\|\zeta_1 \mathcal{L}\varphi\|_s + \|\zeta_1\varphi\|_0\},\quad \varphi \in C_0^\infty(U) \tag{2.15}$$

for all integer s, and all pairs $\zeta, \zeta_1 \in C_0^\infty(U)$ so that $\zeta_1 = 1$ on the support of ζ. The hypoellipticity follows from (2.15). We shall not repeat the arguments leading to (2.15) and (2.14) since at that stage there are no new ideas.

§ 3. Graded and free Lie algebras

In order to formulate a converse to Theorem 1′, and also because of the basic role they will play in what follows, we shall discuss certain particular classes of Lie algebras. Firstly, the Lie algebra $\mathfrak{G}$ is said to be *nilpotent* of *step r*, if $\mathfrak{G}^{(r+1)} = (0)$, where $\mathfrak{G}^{(k)}$ is defined inductively $\mathfrak{G}^{(1)} = \mathfrak{G}$, $\mathfrak{G}^{(k)} = [\mathfrak{G}^{(k-1)}, \mathfrak{G}]$. If $\mathfrak{G} = \mathfrak{N}$ is nilpotent it is well known that the exponential map $\exp \mathfrak{N} \to N$ is a diffeomorphism of $\mathfrak{N}$ onto the corresponding simply connected Lie group N. (See [19].)

A Lie algebra $\mathfrak{N}$ is said to be *graded* if it has a direct sum decomposition $\mathfrak{N} = \sum_{j=1}^r \oplus V^j$, with the property that $[V^j, V^k] \subset V^{k+j}$, if $k+j \leqslant r$ and $[V^k, V^j] = 0$ if $k+j > r$. Observe that a graded algebra is always nilpotent (and of step r). Graded Lie algebras (and in particular the examples below) will be our basic object to study.

Example 1. A *Heisenberg algebra* $\mathfrak{N}$ is a two step graded Lie algebra $\mathfrak{N} = V^1 \oplus V^2$ with the property that $\dim V^2 = 1$, and such that if λ^* is any non-zero linear functional on V^2, the bilinear mapping $(X, Y) \to \lambda^*[X, Y]$ defined on $V^1 \times V^1$ is non-degenerate.

Example 2. A *stratified algebra* (using the terminology of Folland [6]) is any graded Lie algebra with the additional property that V^1 generates the whole Lie algebra. We shall also refer to this kind of algebra as a *stratified algebra of type* I.

Example 3. *Stratified algebra of type* II. This is the graded Lie algebra $\mathfrak{N} = V^1 \oplus V^2 \oplus \dots V^r$, so that there exists a $Y_0 \in V^2$, so that V^1 and Y_0 generate the Lie algebra. This is the kind of Lie algebra that must be used in dealing with the general Hörmander-type second order hypoelliptic operators, as in § 18 below.

Example 4. *Free algebras* $\mathfrak{N}_{F,r}$. For each n and $r \geq 1$ this is the algebra having n generators and r steps, but otherwise as few relations among the commutators as possible. To define this algebra consider first the (infinite-dimensional) free Lie algebra $\mathfrak{G}_F$ on n generators $Y_1, \dots, Y_n$. (Cf. Jacobson [11], Chapter V, § 4.) Roughly speaking, $\mathfrak{G}_F$ is generated by $\{Y_j\}$ with the only relations among the commutators being those forced by anticommutativity and the Jacobi identity. For $r \geq 1$, let $\mathfrak{N}_{F,r} = \mathfrak{G}_F/\mathfrak{G}_{F(r+1)}$. Then $\mathfrak{N}_{F,r}$ is nilpotent of step r and it has the universal property that if $\mathfrak{M}$ is any other nilpotent Lie algebra of step r with n generators, there is a surjective homomorphism of $\mathfrak{N}_{F,r}$ onto $\mathfrak{M}$.

$\mathfrak{N}_{F,r}$ is graded, $\mathfrak{N}_{F,r} = \sum_{j=1}^r \oplus V^j$, with V^j being spanned by all commutators of the form

$$[Y_{k_1}[Y_{k_2} \dots [Y_{k_{j-1}}, Y_{k_j}]] \dots].$$

Example 5. *Two-step algebra* $\mathfrak{N}_2$, *associated to a graded algebra* $\mathfrak{N}$. To every graded Lie algebra $\mathfrak{N}$ of step r, $r \geq 2$, we can associate in a canonical way a two-step graded Lie algebra $\mathfrak{N}_2$. Write $\mathfrak{N} = \sum \oplus V^j$ and take $\mathfrak{N}_2 = \mathfrak{N}/\sum_{j>2} \oplus V^j$.

Observe that $\sum_{j>2} \oplus V^j$ is an ideal, and $\mathfrak{N}_2$ is a graded algebra of two steps.

The converse of Theorem 1′ is as follows.

THEOREM 2. *Let* $\mathfrak{N} = \sum_{j=1}^r V^j$ *be a graded Lie algebra of step* r, $r \geq 2$. *Suppose that* V^1 *is spanned by* $Y_1, Y_2 \dots Y_n$. *Let*

$$\mathcal{L} = \sum_{j=1}^{n} Y_j^2 + \tfrac{1}{2} \sum_{j,k} c_{jk}[Y_j, Y_k]$$

where $c = a + ib$ *and* a, b *are real skew-symmetric matrices. Assume:*

$$\sup_{\varrho \in R^{\perp},\, \|\varrho\|_1 \leq 1} |\mathrm{tr}(b\varrho)| \geq 1, \tag{3.1}$$

and

$$\mathfrak{N}_2' = \mathfrak{N}_2 / (\sum_{j,k} a_{jk}[Y_j, Y_k]) \tag{3.2}$$

is not a Heisenberg algebra. Then $\mathcal{L}$ *is not hypoelliptic.*

Remark. The situation when $\mathfrak{N}_2'$ is a Heisenberg algebra is more complicated and still somewhat obscure. Thus the proof of Theorem 2 shows that $\mathcal{L}$ is not hypoelliptic when $\sup_{\varrho\in R^\perp, \|\varrho\|_1\leqslant 1}|\operatorname{tr}(b\varrho)|=1$ even if $\mathfrak{N}_2'$ is a Heisenberg algebra. Moreover, if $\mathfrak{N}$ is itself a Heisenberg algebra, it is shown in Folland-Stein [8] and Boutet de Monvel-Trèves [3] that $\mathcal{L}$ may be hypoelliptic even if $\sup_{\varrho\in R^\perp, \|\varrho\|_1\leqslant 1}|\operatorname{tr} b\varrho|>1$.

§ 4. Harmonic analysis on N and the proof of Theorem 2

Let $\mathcal{L}$ be the differential operator defined in Theorem 2. We shall use the representation theory of the nilpotent group N in order to find a function φ such that $\mathcal{L}\varphi=0$, but $\varphi\notin C^\infty$.

Recall that a unitary representation of a Lie group N is a continuous homomorphism $\pi: N\to \text{Unit } \mathcal{H}_\pi$, the unitary operators on a Hilbert space $\mathcal{H}_\pi$. By differentiation one obtains a representation $d\pi$ of $\mathfrak{N}$ as skew Hermitian operators on $\mathcal{H}_\pi$. The motivation for considering unitary representations to investigate the hypoellipticity of $\mathcal{L}$ comes from the Plancherel formula. Let $\hat{N}$ be the set of all unitary irreducible representations of N and $d\mu$ the Plancherel measure on $\hat{N}$. For $f\in C_0^\infty(N)$ and $\pi\in\hat{N}$ let $\pi(f)$ be the operator $\int_N \pi(x)f(x)\,dx$. Then $f\to\pi(f)$ extends to an isomorphism of the square integrable functions on N with the space of all $\pi(f)$ with inner product $\int_{\hat{N}}\operatorname{tr}(\pi(f)\pi(g)^*)\,d\mu(\pi)$. (See e.g. Pukanszky [19] for the representation theory of nilpotent Lie groups.)

This indicates that it might be possible to decompose $\mathcal{L}$ as a direct integral

$$\int_{\hat{N}} d\pi(\mathcal{L})\,d\mu(\pi).$$

If $\mathcal{L}$ were hypoelliptic, we might try to invert $\mathcal{L}$ by inverting each operator $d\pi(\mathcal{L})$. An obstruction to such an inversion would be a zero eigenvector of $d\pi(\mathcal{L})$. Thus we are led to look for those π for which $d\pi(\mathcal{L})v=0$ for some $0\neq v\in\mathcal{H}_\pi$.

Fortunately, we shall not have to make use of any deep results in representation theory; we shall be able to construct explicitly all representations which will be needed. In effect these representations will all factor through the quotient algebra $\mathfrak{N}_2'$.

Proof of the theorem. We shall prove the theorem first in the case when $\mathfrak{N}$ is a two-step graded Lie algebra and $\sum a_{jk}[Y_j, Y_k]=0$. In that case $\mathfrak{N}=\mathfrak{N}_2'$, and our assumption is that $\mathfrak{N}$ is not a Heisenberg algebra. We can therefore write $\mathfrak{N}=V^1\oplus V^2$ with $[V^1, V^1]=V^2$. (If $[V^1, V^1]$ is strictly contained in V^2 then $Y_1, \ldots, Y_n$ do not generate $\mathfrak{N}$ and then clearly $\mathcal{L}$ is not hypoelliptic. Hence there is nothing to prove.)

There are now two cases:

$$\text{(i) } \dim(V^2)\geqslant 2;\quad \text{(ii) } \dim(V^2)=1.$$

Since $\sum a_{jk}[Y_j, Y_k]=0$, R is the linear space of relations in V^2, i.e. if $Y_1, ..., Y_n$ form a basis of V^1, $R=\{(s_{jk}): \sum s_{jk}[Y_j, Y_k]=0\}$. Thus the condition (i) implies that $\dim(R^\perp)=\dim V^2\geqslant 2$. So then by continuity we can choose a $\varrho_0\in R^\perp$ with $\varrho_0\neq 0$ so that $\operatorname{tr}(b\varrho_0)=0$. Hence by continuity again we can choose a $\varrho\in R^\perp$ so that $-\operatorname{tr}(b\varrho)=\|\varrho\|_1=1$ (since $\sup_{\varrho\in R^\perp, \|\varrho\|_1\leqslant 1}|\operatorname{tr}(b\varrho)|\geqslant 1$). In case (ii) we have little choice since $R^\perp$ is one-dimensional. We choose ϱ, so that

$$-\operatorname{tr}(\varrho b)\geqslant\|\varrho\|_1=1.$$

Now each $\varrho=(\varrho_{jk})\in R^\perp$ defines for us a linear functional ϱ^* on V^2 by the equations $\varrho^*([Y_j, Y_k])=\varrho_{jk}$. Since $\varrho\in R^\perp$ implies $\sum\varrho_{jk}s_{jk}=0$ whenever $\sum s_{jk}[Y_j, Y_k]=0$, the linear functional is well defined. In case (ii), hypothesis (3.2) of the theorem assures that the bilinear form ϱ defined on $V^1\times V^1$ by $\varrho(Y, Y')=\varrho^*[Y, Y']$ is degenerate. With the element $\varrho\in R^\perp$ chosen as above we proceed with the proof.

Let $\pm i\tilde{\varrho}_j$, $j=1, 2, ..., l$, $2l\leqslant n$, be the non-zero eigenvalues of ϱ. By an orthogonal change of basis we can find $\{Y'_j\}$ satisfying, for $j<k$

$$\varrho^*((Y'_j, Y'_k])=\begin{cases}0 & \text{if } j \text{ is even or } k\neq j+1\\ \tilde{\varrho}_j & \text{if } j \text{ odd, } k=j+1\\ 0 & j\geqslant 2l+1.\end{cases}\tag{4.1}$$

Now for every λ, $0<\lambda<\infty$ we construct a unitary representation π_λ of N on $\mathcal{H}_\lambda=L^2(\mathbf{R}^l)$. π_λ is defined by

$$\pi_\lambda(\exp tY'_j)f(t_1, t_2, ..., t_j, ..., t_l)=f(t_1, t_2, ..., t_k+t, ..., t_l)\quad\text{if } j=2k-1,\ k\leqslant l$$

$$\pi_\lambda(\exp tY'_j)f=e^{it\lambda\tilde{\varrho}_k t_k}f \text{ if } j=2k,\ k\leqslant l$$

$$\pi_\lambda(\exp tY'_j)f=e^{itc(\lambda)}f,\quad\text{where } c(\lambda) \text{ is a constant, if } j=2l+1$$

$$\pi_\lambda(\exp tY'_j)f=f\quad\text{if } j>2l+1.$$

In case (i) we shall take $c(\lambda)=0$ for all λ. We leave the determination of $c(\lambda)$ in case (ii) for later.

Differentiation of the above gives the corresponding representation $d\pi_\lambda$ on $\mathfrak{N}$ defined on the dense space of smooth functions S.

$$d\pi_\lambda(Y'_j)f=\begin{cases}\dfrac{\partial f}{\partial t_k} & j=2k-1,\quad k\leqslant l\\ i\lambda\tilde{\varrho}_k t_k f & j=2k,\quad k\leqslant l\\ ic(\lambda)f & j=2l+1\\ 0 & j>2l+1.\end{cases}\tag{4.2}$$

If $j=2k-1$ then

$$d\pi_\lambda([Y'_j, Y'_{j+1}]) = -d\pi_\lambda([Y'_{j+1}, Y'_j]) = i\lambda\tilde{\varrho}_k.$$

$d\pi_\lambda$ vanishes on all other commutators. From this one obtains

$$d\pi_\lambda([Y_j, Y_k]) = i\lambda\varrho_{jk}. \tag{4.3}$$

From (4.2) and (4.3) we obtain

$$d\pi_\lambda(\mathcal{L}) = -(c(\lambda))^2 + \sum_{k=1}^{l}\left(\frac{\partial^2}{\partial t_k^2} - \lambda^2\varrho_k^2 t_k^2\right) + \frac{i\lambda}{2}\sum_{j,k}(a_{jk}+ib_{jk})\varrho_{kj}.$$

We make the change of variables

$$\eta_k = \sqrt{\lambda|\tilde{\varrho}_k|}\, t_k.$$

Noting also that $a\in R^\perp$, so that $\operatorname{tr}(a\varrho)=0$, we obtain

$$d\pi_\lambda(\mathcal{L}) = \sum_{k=1}^{l}\lambda|\tilde{\varrho}_k|\left(\frac{\partial^2}{\partial\eta_k^2} - \eta_k^2\right) + \frac{i\lambda}{2}\sum_{j,k} ib_{jk}\varrho_{jk} - (c(\lambda))^2. \tag{4.4}$$

Fortunately, the eigenfunctions for the harmonic oscillator $\partial^2/\partial\eta_k^2 - \eta_k^2$ are well known. The Hermite functions

$$H_N(\eta_k) = e^{\eta_k^2/2}\frac{d^N}{d\eta_k^N}(e^{-\eta_k^2})$$

satisfy

$$\left(\frac{\partial^2}{\partial\eta_k^2} - \eta_k^2\right)H_N = -(2N+1)H_N, \; N=0,1,2,\ldots.$$

Furthermore, the collection of products

$$H_{N_1,N_2,\ldots,N_l} = H_{N_1}(\eta_1)H_{N_2}(\eta_2)\ldots H_{N_l}(\eta_l)$$

forms an orthogonal basis of $L^2(\mathbf{R}^l)$, so that the eigenvalues of

$$\sum_{k\leq l}\lambda|\tilde{\varrho}_k|\left(\frac{\partial^2}{\partial\eta_k^2} - \eta_k^2\right) \quad \text{are} \quad -\lambda\sum_{k\leq l}|\tilde{\varrho}_k|(2N_k+1), \; N_k = 0,1,2,\ldots.$$

In particular, take $N_i=0$, all i, and put $H=H_{0\ldots0}$, then

$$\sum\lambda|\tilde{\varrho}_k|\left(\frac{\partial^2}{\partial\eta_k^2} - \eta_k^2\right)H = -\lambda\sum_k|\tilde{\varrho}_k|H = -\frac{\lambda}{2}\|\varrho\|_1 H.$$

Hence from (4.4) we have

$$-\frac{\lambda}{2}\|\varrho\|_1-\frac{\lambda}{2}\operatorname{tr}(b\varrho)-(c(\lambda))^2$$

as the eigenvalue of $d\pi_\lambda(\mathcal{L})$ on H.

Now define

$$c(\lambda)=\left(-\frac{\lambda}{2}(\operatorname{tr}(b\varrho)-\|\varrho\|_1)\right)^{1/2}.$$

The previous discussion insured that the quantity inside the parentheses is non-negative. We have therefore proved the following.

(4.5) LEMMA. *The Hermite function $H=H_{0,0,\ldots,0}$ satisfies $d\pi_\lambda(\mathcal{L})H=0$ for all $\lambda>0$.*

Next we show

(4.6) LEMMA. *Let $\varphi(x)=\int_1^\infty\langle\pi_\lambda(x)H,H\rangle\lambda^{-4}d\lambda$, $x\in N$, where $\langle,\rangle$ is the usual inner product on $\mathcal{H}=L^2(\mathbf{R}^l)$. Then $\mathcal{L}\varphi=0$.*

Proof. Note first that if $Y\in\mathfrak{N}$, and $x\in N$ then

$$Y(\pi_\lambda(x)H)=\frac{d}{dt}\pi_\lambda(x\exp tY)H|_{t=0}=\pi_\lambda(x)\,d\pi_\lambda(Y)H. \tag{4.7}$$

By (4.2) and (4.3) it follows that for any $X_1, X_2, \ldots, X_k\in\mathfrak{N}$, and for fixed $x\in N$, the function $(X_1X_2\ldots X_k\pi_\lambda(x)H)\bar{H}$ is absolutely integrable on $\mathbf{R}^l$. Therefore,

$$(X_1X_2\ldots X_k)\langle\pi_\lambda(x)H,H\rangle=\langle\pi_\lambda(x)d\pi_\lambda(X_1X_2\ldots X_k)H,H\rangle.$$

In particular,

$$\mathcal{L}\langle\pi_\lambda(x)H,H\rangle=\langle\pi_\lambda(x)d\pi_\lambda(\mathcal{L})H,H\rangle=0.$$

It will follow that $\mathcal{L}\varphi=0$ if differentiation under the integral sign can be justified. However, by (4.7) and (4.2), since $\mathcal{L}$ is an operator of order 2, $\mathcal{L}\langle\pi_\lambda(x)H,H\rangle$ grows at most as λ^2 for x fixed. Hence $\mathcal{L}\langle\pi_\lambda(x)H,H\rangle\lambda^{-4}$ is absolutely integrable on $[1,\infty)$. Therefore,

$$\mathcal{L}\int_1^\infty\langle\pi_\lambda(x)H,H\rangle\lambda^{-4}d\lambda=\int_1^\infty\mathcal{L}\langle\pi_\lambda(x)H,H\rangle\lambda^{-4}d\lambda=0.$$

To complete the proof of Theorem 2 (in the case $\mathfrak{N}=\mathfrak{N}_2'$), it suffices to show that $\varphi\notin C^\infty(N)$. We shall prove this by showing that the restriction of φ to $\exp(\mathbf{R}[Y_j,Y_k])$ is not smooth for some j,k. In fact, choose j,k so that $\varrho_{jk}\neq0$, and note that by (4.3), $\pi_\lambda(\exp t(Y_j,Y_k])=e^{it\lambda\varrho_{jk}}$. Put $c=\varrho_{jk}$ for convenience and $\varphi(t)=\varphi(\exp t[Y_j,Y_k])$; it suffices to show

$$\varphi(t)=\int_1^\infty e^{ic\lambda t}\lambda^{-4}d\lambda\notin C^\infty(\mathbf{R}).$$

To see this, note that since φ is obviously twice differentiable, we are reduced to showing

$$\int_1^\infty e^{ic\lambda t}\lambda^{-2}d\lambda \notin C^\infty(\mathbf{R}).$$

This may be proved by writing

$$\int_1^\infty e^{ic\lambda t}\lambda^{-2}d\lambda = \int_1^\infty \cos c\lambda t\,\lambda^{-2}d\lambda + i\int_1^\infty \sin c\lambda t\,\lambda^{-2}d\lambda = I_1(t)+I_2(t)$$

and showing that the derivative of $I_2(t)$ is unbounded as $t\to 0$. Details are left to the reader.

This finishes the proof when $\mathfrak{N}=\mathfrak{N}_2$. In the general case we adopt the notation (and implicitly the point of view) we shall also use later.

Let $\tilde{\mathfrak{N}}=\sum_{i=1}^r \oplus \tilde{V}^r$ be a graded Lie algebra with $\tilde{Y}_1, \tilde{Y}_2 \dots \tilde{Y}_n$ spanning $\tilde{V}^1$, and let $h: \tilde{\mathfrak{N}}\to\mathfrak{N}_2' = V^1+V^2 = (\tilde{\mathfrak{N}}/\sum_{j>2}\oplus \tilde{V}^j)/(\sum_{j,k}a_{jk}[\tilde{Y}_j, \tilde{Y}_k])$. Observe that $h(\tilde{V}^i)=V^i$, $i=1, 2$, and that h is an isomorphism from $\tilde{V}^1$ to V^1 and a surjection of $\tilde{V}^2$ onto V^2 with kernel $\sum_{j,k}a_{jk}[\tilde{Y}_j, \tilde{Y}_k]$.

Let $\tilde{N}$ and N denote the simply connected Lie groups corresponding to $\mathfrak{N}$ and $\mathfrak{N}_2'$ respectively, and N_0 that of the Lie subalgebra of $\tilde{\mathfrak{N}}$ spanned by $\sum_{i>2}\oplus\tilde{V}^i$ and $\sum_{j,k}a_{jk}[\tilde{Y}_j, \tilde{Y}_k]$. Then N_0 is a normal subgroup of $\tilde{N}$, and $\tilde{N}/N_0=N$. Denote also by h the canonical homomorphism $h: \tilde{N}\to\tilde{N}/N_0=N$. $Y_j=h(\tilde{Y}_j)$, $j=1, \dots, n$, is a basis for V^1. We need to show that the operator $\tilde{\mathcal{L}}=\sum \tilde{Y}_i^2+\frac{1}{2}\sum c_{jk}[\tilde{Y}_j, \tilde{Y}_k]$ is not hypoelliptic under hypotheses (3.1) and (3.2). Now we have already seen that the corresponding operator defined on the group N is not hypoelliptic; in fact we showed the existence of a φ, so that $\varphi\notin C^\infty$ but $\mathcal{L}(\varphi)=0$. Let now $\tilde{\varphi}(\tilde{x})=\varphi(h(\tilde{x}))$, $\tilde{x}\in\tilde{N}$. Then since $\tilde{Y}_j(\tilde{\varphi})(\tilde{x})=(Y_j\varphi)(h(\tilde{x}))$, and $\tilde{Y}_j\tilde{Y}_k(\tilde{\varphi})(\tilde{x})=(Y_jY_k\varphi)(h\tilde{x})$, it follows that $\tilde{\mathcal{L}}(\tilde{\varphi})=0$; but $\tilde{\varphi}$ is not in $C^\infty(\tilde{N})$ and so the theorem is completely proved.

§ 5. Dilations and homogeneity on groups

A *family of dilations* on a nilpotent Lie algebra $\mathfrak{N}$ is a one-parameter group $\{\delta_t\}_{t\in\mathbf{R}^+}$ of automorphisms of $\mathfrak{N}$ determined by $\delta_t(Y_j)=t^{a_j}Y_j$, where $\{Y_j\}_{1\leq j\leq m}$ is a linear basis for $\mathfrak{N}$ and $\{a_j\}_{1\leq j\leq m}$ is a set of positive real numbers. By the exponential map δ_t lifts to a one-parameter group of automorphisms of N, the simply connected nilpotent Lie group corresponding to $\mathfrak{N}$; these automorphisms will again be denoted by $\{\delta_t\}$. N, equipped with these dilations, is then called a *homogeneous group.*

Any graded nilpotent Lie algebra $\mathfrak{N}$ has a natural family of dilations $\{\delta_t\}$. Indeed, if $\mathfrak{N}=\sum_{j=1}^r V^j$, it is not hard to see that the mappings δ_t defined on each V^j by $\delta_t(Y)=t^jY$,

$Y \in V^j$, $t>0$, extend by linearity to automorphisms of $\mathfrak{N}$. In this paper we shall entirely restrict our attention to these dilations on graded Lie algebras.

In this section we review some facts about operators on homogeneous groups, and refer the reader to Knapp-Stein [12], Koranyi-Vagi [16], Folland-Stein [8], and especially Folland [6], for details.

A *homogeneous norm function* on a homogeneous group N is a mapping $x \to |x|$, $x \in N$ satisfying

$$\begin{cases} \text{(i)}\ |x| \geqslant 0 \text{ and } |x|=0 \text{ if and only if } x=0, \\ \text{(ii)}\ x \to |x| \text{ is continuous on } N \text{ and smooth on } N-\{0\}, \\ \text{(iii)}\ |\delta_t(x)| = t|x|. \end{cases} \tag{5.1}$$

Since we shall consider here the case where $\mathfrak{N}=\sum_{j=1}^{r} V^j$ is graded, we may exhibit a particular norm function as follows. Any $x \in N$ has a unique representation

$$x = \exp(Y^1+Y^2+\dots Y^r),\ Y^j \in V^j.$$

Then we may define $|\ |$ by

$$|x| = \left(\sum_{j=1}^{r} \|Y^j\|^{(2r!)/j}\right)^{1/(2r!)}, \tag{5.2}$$

where $\|\ \|$ indicates the Euclidean norm on V^j.

We shall occasionally need the vector space sum of points of N: if $x, y \in N$ we write $x+y$ for $\exp(\log x + \log y)$. Then we have the "triangle inequalities"

$$\begin{gathered} |x+y| \leqslant \gamma(|x|+|y|), \text{ and} \\ |xy| \leqslant \gamma(|x|+|y|) \end{gathered} \tag{5.3}$$

for some constant $\gamma \geqslant 1$. Writing $\|\ \|$ for the Euclidean norm on N there exist positive constants C_1 and C_2 such that for all $|x| \leqslant 1$,

$$C_1\|x\| \leqslant |x| \leqslant C_2\|x\|^{1/r}. \tag{5.4}$$

The *homogeneous dimension* of N is defined by

$$Q = \sum_{j=1}^{r} j(\dim V^j). \tag{5.5}$$

It significance is that if dx is Lebesgue measure on N, then $dx/|x|^Q$ is invariant under the dilations δ_t, $t>0$.

A measurable function f on N will be called *homogeneous of degree* λ if $f\circ\delta_t=t^\lambda f$, all $t>0$. Any function K which is homogeneous of degree $-Q+\alpha$, $0<\alpha$ and smooth away from the origin is locally integrable and thus defines a distribution T. T will be called a homogeneous distribution of type α. Similarly, suppose K is a homogeneous function of degree $-Q$ which is smooth away from the origin and satisfies the mean value property $\int_{a<|x|<b}K(x)\,dx=0$, all a, b, $0<a<b$. If $c\in\mathbf{C}$, then the pair (K, c) defines a distribution T given by

$$T(f)=\lim_{\varepsilon\to 0}\int_{\varepsilon<|x|<\frac{1}{\varepsilon}} K(x)\,f(x)\,dx+cf(0), \tag{5.6}$$

$f\in C_0^\infty(N)$. T will be called a homogeneous distribution of type 0. With a slight abuse of notation we shall also denote a homogeneous distribution of type 0 by K.

A differential operator D or N is called *homogeneous* of *degree* λ if $D(f\circ\delta_t)=t^\lambda(Df)\circ\delta_t$, all $t>0$. If f is a homogeneous function of degree α and D is homogeneous of degree λ, then Df is a homogeneous function of degree $\alpha-\lambda$. If τ is a homogeneous distribution of type $\alpha\geqslant\lambda$, then $D\tau$ is of type α-λ.

Recall that if f and g are functions on N their convolution $f*g$ is defined by

$$f*g(y)=\int_N f(x)\,g(x^{-1}y)\,dx=\int f(yx^{-1})\,g(x)\,dx. \tag{5.7}$$

If $f\in C_0^\infty$ and τ is a distribution we may define $f*\tau$ and $\tau*f$ as C^∞ functions by $(f*\tau)(x)=\tau(f^x)$, where $f^x(y)=f(xy^{-1})$ and $(\tau*f)(x)=\tau(f_x)$, where $f_x(y)=f(y^{-1}x)$.

If τ is actually a function these definitions agree with the usual notion of convolution given by (5.7). If D is a left-invariant differential operator the reduction to the case $D=Y\in\mathfrak{N}$ shows

$$D(f*\tau)=f*D\tau, \quad\text{and}\quad D(\tau*f)=\tau*(Df). \tag{5.8}$$

We shall need to discuss next the existence of fundamental solutions for a certain class of left-invariant differential operators on a graded Lie group. For some of the applications below we shall have to deal with systems of such differential operators. For this purpose we shall assume that our functions take their values in a finite-dimensional vector space W over $\mathbf{C}$. The coefficients of the differential operator $\mathcal{L}$ in question, as well as the fundamental solution K, will then take their values in the space of linear transformations of W to itself. Thus the notions of homogeneity defined above for scalar functions and operators may be extended to this case by requiring the appropriate homogeneity of each component or matrix entry.

Now let $D=(D_{jk})$ be a homogeneous differential operator of degree λ and $K=(K_{jk})$

a homogeneous distribution of type α. We denote by DK the matrix of scalar distributions $(DK)_{jk} = \sum_l D_{jl} K_{lk}$. Each $(DK)_{jk}$ is a scalar distribution of type $\alpha-\lambda$, by the above remarks.

(5.9) LEMMA. *Let D, K be as above. Suppose that D is left-invariant. Let T be the operator given on smooth compactly supported functions f by $T(f) = f * K$. Then*

$$D(T(f)) = D(f * K) = f * (DK). \tag{5.10}$$

Proof. The content of (5.10) is that convolution by the matrix product DK corresponds to the composition of the operators D and T when D is left-invariant. This follows from the corresponding statement (5.8) in the scalar case. Since every component function of DK is smooth away from 0 and homogeneous of degree $(-Q+\alpha)-\lambda$, the proof of the last statement reduces to showing that if $\alpha=\lambda$, the components have mean value 0. This is immediate from the scalar case. (See e.g. Folland [6], Proposition 1.8.) q.e.d.

Now if a fixed (positive definite) inner product, $(,)_0$, has been given in W, we may define $(f, g) = \int_N (f, g)_0(x)\,dx$ whenever $f, g \in C_0^\infty(N, W)$. We write $\mathcal{L}^*$ for the formal adjoint of $\mathcal{L}$, i.e. $(\mathcal{L}f, g) = (f, \mathcal{L}^* g)$ whenever f, g are smooth and have compact support.

The following theorem concerning the existence of fundamental solutions for left-invariant differential operators will be crucial in our construction of parametrices.

PROPOSITION A.(1) *Let $\mathcal{L}$ be a left-invariant hypoelliptic differential operator on N such that the formal adjoint $\mathcal{L}^*$ is also hypoelliptic. If $\mathcal{L}$ is homogeneous of degree α, $0<\alpha<Q$, then there is a unique homogeneous distribution k of type α such that for all $f \in C_0^\infty(N)$,*

$$\mathcal{L}(f * k) = (\mathcal{L}f) * k = f. \tag{5.11}$$

We shall now state several results which are known in the scalar case; the extension to the vector-valued case is immediate.

A fundamental result on convolution by homogeneous distributions is the following.

PROPOSITION B. *Let τ be a homogeneous distribution of type α, $0 \leq \alpha < Q$. If $\alpha = 0$ then convolution by τ extends from $C_0^\infty(N)$ to a bounded mapping on $L^p(N)$, $1<p<\infty$. If $\alpha>0$ convolution by τ extends to a bounded map from L^p to L^q, where $1/q = (1/p) - (\alpha/Q)$ provided $1<p<Q/\alpha$.*

In defining the convolution of matrix-valued functions K and L we must take into account the fact that in general $L(x)K(x) \neq K(x)L(x)$. We put

(1) See Folland [6], Corollary 2.8. A similar result was also obtained both by R. Strichartz and the second author (unpublished). The result in Folland [6] is stated in the scalar case (dim $W = 1$); the same argument holds as well in the general case.

$$(K * L)(x) = \int L(y) K(xy^{-1}) dy.$$

PROPOSITION C. *Let K_α, K_β be homogeneous distributions of type $\alpha, \beta \geqslant 0$, respectively with $Q > \alpha+\beta > 0$. Then $K_\alpha * K_\beta$ exists, and is a homogeneous distribution of type $\alpha+\beta$. Furthermore, the associative law*

$$(f * K_\alpha) * K_\beta = f * (K_\alpha * K_\beta) \tag{5.12}$$

holds for all $f \in L^p$, if $p < Q/(\alpha+\beta)$.

§ 6. Smoothly varying families of fundamental solutions

Let $\mathfrak{N} = \sum_{j=1}^{r} V^j$ be a graded Lie algebra and let

$$\mathcal{L} = \sum_{j=1}^{n} Y_j^2 + \tfrac{1}{2} \sum_{j,k} c_{jk} [Y_j, Y_k] \tag{6.1}$$

where Y_i is a basis of V^1, V^1 is assumed to generate $\mathfrak{N}$, and $(c_{jk}) = c = a + ib$, where a and b are real skew-symmetric matrices.

By Theorem 1′ we know that whenever c_{jk} lies in a certain (open) subset of the $n(n-1)$-dimensional parameter space $\mathcal{L}$ is hypoelliptic. Proposition A in § 5 then tells us that $\mathcal{L}$ has associated to it a unique fundamental solution, which is a distribution of type 2. We shall be concerned with proving that this fundamental solution varies smoothly with the (c_{ij}).

The considerations of the $\bar\partial_b$ problem in § 19 below will require that we prove our results for systems. As we already did in § 5 we shall assume that our functions take values in a finite-dimensional vector space W. The operator $\mathcal{L}$ in (6.1) is then replaced by $(\sum_{j=1}^{n} Y_j^2) I + \frac{1}{2} \sum_{j,k} c_{jk} [Y_j, Y_k]$ where each c_{jk} is a linear transformation of W to itself. However now the necessary and sufficient conditions on the c_{jk} that $\mathcal{L}$ is hypoelliptic are no longer as simple as those given in Theorems 1′ and 2. We shall assume nevertheless that for some fixed (c_{jk}) the two a-priori inequalities hold:

$$\sum_{j=1}^{n} \| Y_j f \|^2 \leqslant C |(\mathcal{L} f, f)|, \quad f \in C_0^\infty(N)$$

and

$$\sum_{j=1}^{n} \| Y_j f \|^2 \leqslant C |(\mathcal{L}^* f, f)|, \quad f \in C_0^\infty(N).$$

We now define $\mathcal{L}_\gamma$ by

$$\mathcal{L}_\gamma = \mathcal{L} + \sum \gamma_{jk} [Y_j, Y_k].$$

Then for γ ranging in some open subset Ω of the parameter space (of real dimension $= n(n-1)\times(\dim W)^2$), and in particular for γ sufficiently small, the operators $\mathcal{L}_\gamma$ and $\mathcal{L}_\gamma^*$ satisfy $\sum_{j=1}^n \|Y_j f\|^2 \leqslant C_\gamma |(\mathcal{L}_\gamma f, f)|$ and $\sum_{j=1}^n \|Y_j f\|^2 \leqslant C_\gamma |(\mathcal{L}_\gamma^* f, f)|$, for $f\in C_0^\infty(N)$. Thus by the reasoning already quoted in § 2, both $\mathcal{L}_\gamma$ and $\mathcal{L}_\gamma^*$ are hypoelliptic. Now by Proposition A in § 5 it follows that for each $\gamma\in\Omega$, there exists a unique homogeneous distribution k_γ of type 2 on N, so that $\mathcal{L}_\gamma(k_\gamma)=I\delta$, and such that $\mathcal{L}_\gamma(f*k_\gamma)=(\mathcal{L}_\gamma f)*k_\gamma=f$ for all $f\in C_0^\infty(N)$.

Our result on the smooth dependence of k_γ on γ may be stated as follows.

THEOREM 3. *The function $(\gamma, x)\to k_\gamma(x)$ is (jointly) C^∞ on the set $\Omega\times(N-\{0\})$.*

Since each function $k_\gamma(x)$, $\gamma\in\Omega$, is homogeneous in x of the same degree, it suffices to prove smoothness on the set $\Omega\times\{x\in N: 1\leqslant|x|\leqslant 2\}$, where $|\ |$ is a homogeneous norm function. For this purpose we define $\mathfrak{C}^\infty$ to be the space consisting of all complex valued functions which are smooth on $\{x\in N: 1<|x|<2\}$ and which, together with all their derivatives, are continuous on the closure $\{x\in N: 1\leqslant|x|\leqslant 2\}$. On $\mathfrak{C}^\infty$ we define a countable collection of seminorms $\{\|\ \|_\alpha\}$, $\alpha=(\alpha_1,\dots,\alpha_n)$ a multi-index of non-negative integers by

$$\|f\|_\alpha=\sup_{1\leqslant|x|\leqslant 2}\left|\left(\frac{\partial}{\partial x}\right)^\alpha f(x)\right|.$$

The space $\mathfrak{C}^\infty$ is complete with respect to these seminorms.

We shall say that a mapping $f\colon \Omega\to\mathfrak{C}^\infty$ is *bounded* if the positive functions $\gamma\to\|f(\gamma)\|_\alpha$ are bounded for each seminorm. Our proof will proceed in two steps.

Step 1. The mapping $\gamma\to k_\gamma$ of Ω to $\mathfrak{C}^\infty$ is bounded on any compact subset of Ω.

Step 2. For each $x\in N-\{0\}$ the function $\gamma\to k_\gamma(x)$ is C^∞, and its partial derivatives with respect to γ are bounded on compact subsets of $\Omega\times\{x: 1\leqslant|x|\leqslant 2\}$.

Now Theorem 3 is proved from steps 1 and 2 by the following lemma.

(6.2) LEMMA. *Suppose a complex valued function $(\gamma, x)\to F(\gamma, x)$ is defined on an open subset $U\subset\mathbf{R}^{n_1}\times\mathbf{R}^{n_2}$ and satisfies the following properties:*

(i) *F is C^∞ in both variables x and γ, separately.*

(ii) *All partial derivatives of F with respect to either x or γ are bounded on compact subsets of U.*

Then F is jointly C^∞ on U.

This result is certainly not new, but, lacking an explicit reference, we give a proof. Since the conclusion is local we may assume (after multiplying by suitable C^∞ functions of compact support) that F also has compact support. Let $\hat F(\gamma', x')$ be the Fourier transform of F. By our assumptions both $|\gamma'|^k\hat F(\gamma', x')$ and $|x'|^k\hat F(\gamma', x')$ belong to L^2 for any

$k>0$. Therefore $(|x'|^2+|y|^2)^{k/2}\hat{F}(y', x')$ is in L^2 for all $k>0$, so $F\in C^\infty$ by the Fourier inversion formula.

We shall need some preliminaries before proceeding to the proof of steps 1 and 2.

For any $\gamma\in\Omega$ write T_γ for the mapping corresponding to the kernel k_γ.

(6.3) LEMMA. *For any* $\gamma_1, \gamma_2\in\Omega$,

$$T_{\gamma_1}-T_{\gamma_2}=T_{\gamma_2}(\mathcal{L}_{\gamma_2}-\mathcal{L}_{\gamma_1})\,T_{\gamma_1}.$$

Proof. Recall that this equality is with reference to smooth functions f of compact support. Evaluating the right hand gives

$$T_{\gamma_2}(\mathcal{L}_{\gamma_2}(T_{\gamma_1}f))-T_{\gamma_2}(\mathcal{L}_{\gamma_1}(T_{\gamma_1}f))=T_{\gamma_1}f-T_{\gamma_2}f$$

by identity (5.11) which proves the lemma.

It will be important in what follows to note that since $\mathcal{L}_{\gamma_2}-\mathcal{L}_{\gamma_1}$ is left-invariant, by Lemma 5.9

$$(\mathcal{L}_{\gamma_2}-\mathcal{L}_{\gamma_1})\,(T_{\gamma_1}f)=(\mathcal{L}_{\gamma_2}-\mathcal{L}_{\gamma_1})\,(T_{\gamma_1})\,f; \tag{6.4}$$

here $(\mathcal{L}_{\gamma_2}-\mathcal{L}_{\gamma_1})\,T_{\gamma_1}$ is the mapping given by convolution by the kernel $(\mathcal{L}_{\gamma_2}-\mathcal{L}_{\gamma_1})\,k_{\gamma_1}$, which is of type 0.

(6.5) LEMMA. *Given* $\gamma_0\in\Omega$, *there exists* $\varepsilon>0$ *such that if* $|\gamma-\gamma_0|<\varepsilon$ *the operator*

$$E_\gamma=(\mathcal{L}_{\gamma_0}-\mathcal{L}_\gamma)\,T_{\gamma_0}$$

satisfies

$$\|E_\gamma f\|_{L^p}\leqslant\tfrac{1}{2}\|f\|_{L^p}. \tag{6.6}$$

Proof. $T_{\gamma_0}(f)=f*k_{\gamma_0}$, k_{γ_0} a homogeneous distribution of type 2. Thus each of the operators $Y_jY_kT_{\gamma_0}$ and $[Y_j, Y_k]\,T_{\gamma_0}$ is given by convolution with a homogeneous distribution of type 0, and is therefore bounded on L^p by Proposition B of § 5. This proves (6.6) for ε sufficiently small.

The following is a key point in the proof of step 1.

(6.7) LEMMA. *The estimates*

$$\|T_\gamma(f)\|_{L^2}\leqslant C_\gamma\|f\|_{L^p} \tag{6.8}$$

hold with $\frac{1}{2}=(1/p)-2/Q$, *and* C_γ *bounded on any compact subset of* Ω.

Proof. It suffices to show that given any $\gamma_0\in\Omega$ there is an ε neighborhood of γ_0 so that (6.8) holds for all $|\gamma-\gamma_0|<\varepsilon$. Let C_0 be the norm of T_{γ_0} as a bounded operator from

L^p to L^2 (see Proposition B of § 5). We shall prove (6.8) provided $C_\gamma \geq 2C_0$ and ε is chosen to satisfy conditions of Lemma 6.5. Then putting $g=\sum_{n=0}^{\infty} E^n(f)$, $E=(\mathcal{L}_{\gamma_0}-\mathcal{L}_\gamma)T_{\gamma_0}$, we have

$$\|g\|_{L^p} \leq \sum_{l=0}^{\infty} 2^{-l}\|f\|_{L^p}=2\|f\|_{L^p}.$$

However, since $T_\gamma(f)=T_{\gamma_0}(g)$, by Proposition B

$$\|T_\gamma(f)\|_{L^2}=\|T_{\gamma_0}(g)\|_{L^2} \leq C_0\|g\|_{L^p} \leq 2C_0\|f\|_{L^p},$$

and Lemma 6.7 is proved.

(6.9) LEMMA. *There exists $\varepsilon>0$ with the following properties: Whenever ζ and $\zeta_1 \in C_0^\infty$ and $\zeta_1=1$ on the support of ζ and $\gamma \in \Omega$*

(6.10) $$\|\zeta f\|_{L^2_{k+\varepsilon}} \leq C(\|\zeta_1 \mathcal{L}_\gamma(f)\|_{L^2_k}+\|\zeta_1 f\|_{L^2_0}).$$

The constant C depends on ζ, ζ_1, k, and γ, but when the first three are fixed, C remains bounded as γ ranges over compact subsets of Ω.

Proof. (6.10), without estimates for the constant C, is proved e.g. in Kohn [15] as a consequence of the estimate

(6.11) $$\sum_{j=1}^{n}\|X_j f\|^2 \leq C_\gamma(|(\mathcal{L}_\gamma f, f)|+\|f\|^2) \quad f \in C_0^\infty.$$

Moreover, (6.11) is proved in Theorem 1′ in the scalar case, and is what we have assumed in the general case. It is clear that C_γ remains bounded on compact subsets of Ω. Now (6.10) with the additional statement about the dependence of C on γ is a consequence merely of keeping track of the constants in Kohn's argument. We omit the details.

We may now prove step 1. Suppose that g is any complex valued function satisfying

(i) supp $g \subset \{x: 1 \leq |x| \leq 2\}$, and

(ii) $\int |g(x)|^2 dx \leq 1$.

Now let $f=T_\gamma(g)=g * k_\gamma$. Since $\mathcal{L}_\gamma T_\gamma(g)=g$, in the sense of distributions, $\mathcal{L}_\gamma(f)=0$ for $|x|<1$. Also $f=\int g(y)K_\gamma(y^{-1}x)dy$ is C^∞ for $|x| \leq 1$. Now apply the inequality (6.10) with ζ, ζ_1 satisfying

$$\zeta_1=1 \quad \text{for } |x| \leq \tfrac{1}{2} \quad \text{and} \quad =0 \quad \text{for } |x| \geq 1,$$
$$\zeta=1 \quad \text{for } |x| \leq \tfrac{1}{4} \quad \text{and} \quad =0 \quad \text{for } |x| \geq \tfrac{1}{2}.$$

Then

(6.12) $$\|\zeta_1 f\|_{L^2} \leq C\|f\|_{L^2}=C\|k_\gamma(g)\|_{L^2} \leq C_\gamma\|g\|_{L^p} \leq C'_\gamma\|g\|_{L^2}$$

where C_γ remains bounded as γ ranges over compact subsets of Ω. The next to last inequality follows from (6.8) and the last from Hölder's inequality since $p<2$ and g is supported in a compact set. In our case inequality (6.10) becomes

$$\|\zeta f\|_{L^2_{k+s}} \leqslant C_{\zeta,k,\gamma}\|g\|_{L^2}, \text{ any } k. \tag{6.13}$$

An application of Sobolev's Lemma then gives

$$\left|\left(\frac{\partial}{\partial x}\right)^\beta f(0)\right| \leqslant C'_{\beta,\gamma}\|g\|_{L^2}, \text{ any multi-index } \beta.$$

In particular if D is any left-invariant differential operator on N i.e. a polynomial in the Y_j's, then

$$|(Df)(0)| \leqslant C_{D,\gamma}\|g\|_{L^2}. \tag{6.14}$$

However, $f=g*k_\gamma$ so that

$$Df(0)=(g*Dk_\gamma)(0)=\int g(y)\,D(k_\gamma(y^{-1}))\,dy. \tag{6.15}$$

Since g is arbitrary among L^2 functions of norm $\leqslant 1$ supported in $\{x;\,1\leqslant|x|\leqslant 2\}$, the converse of Schwarz's inequality can be applied to (6.14) and (6.15). The result is

$$\left(\int_{1\leqslant|y|\leqslant 2}|Dk_\gamma(y^{-1})|^2\,dy\right)^{1/2} \leqslant C_{D,\gamma}.$$

By homogeneity of the kernel $k_\gamma(y^{-1})$ if $0<\delta<1$, there exists $C'_{D,\gamma}$ such that

$$\left(\int_{1-\delta\leqslant|y|\leqslant 2+\delta}|Dk_\gamma(y^{-1})|^2dy\right)^{1/2} \leqslant C'_{D,\gamma}.$$

Finally another application of Sobolev's Lemma gives

$$\sup_{1\leqslant|x|\leqslant 2}\left|\left(\frac{\partial}{\partial x}\right)^\beta k_\gamma(x)\right| \leqslant C_{\beta,\gamma}$$

with $C_{\beta,\gamma}$ bounded as γ ranges over compact subsets of Ω. step 1 is therefore completely proved.

We shall now prove step 2 i.e. that for x fixed the map $\gamma\to k_\gamma(x)$ is C^∞ with partial derivatives bounded on compact subsets of Ω. We have the identities

$$T_\gamma(I-E)=T_{\gamma_0} \quad \text{with } E=(\mathcal{L}_{\gamma_0}-\mathcal{L}_\gamma)\,T_{\gamma_0},$$

(as operators on smooth functions of compact support), which follow from Lemmas (6.3) and (6.4). With the observation $(I-E)(\sum_{k=0}^{l}E^k)=I-E^{l+1}$ we obtain

$$T_\gamma = T_{\gamma_0}\left(\sum_{k=0}^{l} E^k\right) + T_\gamma E^{l+1}. \tag{6.16}$$

We shall interpret (6.16) as a Taylor expansion in γ with remainder. First write

$$E = (\mathcal{L}_{\gamma_0} - \mathcal{L}_\gamma) T_{\gamma_0} = \sum_j (\gamma^j - \gamma_0^j) E_{j,\gamma_0}, \tag{6.17}$$

where γ^j are the coordinates of γ, and each E_j, γ_0 is an operator of the form $E_{j,\gamma_0}(f) = f * k_{j,\gamma_0}$, with k_{j,γ_0} of the form $Y_l^2(k_\gamma)$ or $[Y_l^2, Y_k](k_\gamma)$. (See (5.10).) Hence each k_{j,γ_0} is a homogeneous distribution of type 0. Also, in view of step 1, each k_{j,γ_0} ranges over a bounded subset of C^∞ as γ_0 ranges over a compact subset of Ω.

For each j, consider the operator $T_{\gamma_0} E_{j,\gamma_0}$. By Proposition C of § 5,

$$T_{\gamma_0} E_{j,\gamma_0}(f) = f * (k_{j,\gamma_0} * k_{\gamma_0}).$$

Furthermore, by examining the proof of Proposition C (see e.g. [6], (1.13)) we may obtain the stronger result that if k_1, k_2 are kernels varying over a bounded subset of C^∞, then then $k_1 * k_2$ varies over a bounded subset of C^∞. Thus the kernels $(\gamma^j - \gamma_0^j) K_{j,\gamma_0} * k_{\gamma_0}$ vary over a bounded subset of C^∞ if γ_0, γ are bounded. Now substituting (6.17) into (6.16) we obtain for any f smooth and compactly supported,

$$T_\gamma(f) = f * \left(\sum_{|\alpha| \leq l} \frac{(\gamma - \gamma_0)^\alpha}{\alpha!} k_{\gamma_0}^{(\alpha)} + R_l(\gamma, \gamma_0) |\gamma - \gamma_0|^{l+1}\right), \tag{6.18}$$

where the $k_{\gamma_0}^{(\alpha)}$ and $R_l(\gamma, \gamma_0)$ are homogeneous distributions of type 2 which vary over a bounded subset of C^∞ as γ and γ_0 vary in a compact subset of Ω. However, by definition, $T_\gamma(f) = f * k_\gamma$. Hence

$$k_\gamma(x) = \sum_{|\alpha| \leq l} \frac{(\gamma - \gamma_0)^\alpha}{\alpha!} k_{\gamma_0}^{(\alpha)}(x) + R_l(\gamma, \gamma_0)(x) |\gamma - \gamma_0|^{l+1} \tag{6.19}$$

for all $x \in N - \{0\}$. For x fixed, $1 \leq |x| \leq 2$, (6.19) provides a Taylor expansion with remainder for the function $\gamma \to k_\gamma(x)$; for each l the coefficients $k_{\gamma_0}^{(\alpha)}(x)$ and remainder $R_l(\gamma, \gamma_0)(x)$ are bounded as γ_0 and γ range over a compact subset of Ω. Thus we may apply the converse of Taylor's theorem with remainder. (See e.g. [1], where a slightly weaker form is given i.e. assuming the coefficients are continuous instead of bounded. An easy modification of their argument proves a similar result in the bounded case.) Thus we conclude that $\gamma \to k_\gamma(x)$ is C^∞ on Ω, and for each α,

$$\left(\frac{\partial}{\partial \gamma}\right)^\alpha k_\gamma(x) = k_\gamma^{(\alpha)}(x).$$

Since $k_\gamma^{(\alpha)}$ is bounded for γ in a compact subset of Ω, so is $(\partial/\partial\gamma)^\alpha k_\gamma(x)$, proving the second claim of step 2. This completes the proof of step 2 and therefore of Theorem 3.

Part II. Extension of the manifold and approximation by a free group

§ 7. Lifting of vector fields to free groups

Let M be a real C^∞ manifold of dimension m, and let $X_1, X_2 \ldots X_n$ be real, smooth vector fields on M such that finitely many commutators of the X_i's span the tangent space at every point. We would like to associate to every point $\xi \in M$ a nilpotent Lie group N_ξ and a local diffeomorphism Θ_ξ identifying a neighborhood of ξ in M with one of the identity in N_ξ. In this local coordinate system the $X_1, X_2, \ldots, X_n$ should be closely approximated, in some sense, by left invariant vector fields $Y_1, \ldots, Y_n$ generating the Lie algebra. This is, roughly speaking, the approach used by Folland and Stein [8] to construct a parametrix for the Laplacian $\Box_b$ of the tangential Cauchy-Riemann operator. For this case the vector fields involved satisfy particularly simple commutation relations, and it is therefore possible to assign the same group N_ξ, the Heisenberg group of appropriate dimension, at each point ξ.

In the general case there seems to be no natural group N having $\dim N = \dim M$. Consider, however, Example (b) of the Introduction, where the X_j are not linearly independent at each point. In that example it is necessary to add an extra variable and thus to lift the original vector fields to a higher dimensional space. The resulting manifold may then be identified with a nilpotent Lie group. With this example in mind, our approach in general will be to lift the X_j to a higher dimensional manifold $\tilde{M}$ in such a way as to eliminate inessential relations among the commutators. We will then assign to every point $\tilde{\xi} \in \tilde{M}$ the free nilpotent group on n generators of step r, where r is sufficiently large. We now proceed to the details of this construction.

We shall refer to the elements in the linear span of

$$\{[X_{i_1}, [X_{i_2}, \ldots, [X_{i_{s-1}}, X_{i_s}]] \ldots]: \ 1 \leqslant i_j \leqslant n\}$$

as the commutators of length s, and to the span of $X_1, X_2 \ldots X_n$ as the commutators of length 1 to avoid having to deal with this special case. If X is any vector field and $\xi \in M$ we write $X|_\xi$ for the restriction of X to the tangent space at ξ. Throughout this paper we shall assume that the commutators of length $\leqslant r$ span the tangent space at every $\xi \in M$.

We now make precise the notion that a set of vector fields $\{W_k\}_{k=1,2,\ldots,n}$, and their commutators of lengths $\leqslant s$ satisfy as few linear relations as possible at a given point ξ. To do this, we compare the W_k and their commutators with left-invariant vector fields on a free nilpotent Lie algebra. Let n_s be the dimension of the free nilpotent Lie algebra $\mathfrak{N}_{F,s}$ of step s on n generators, and let m_s be the dimension of the linear space spanned by all commutators of the W_k of lengths $\leqslant s$ restricted to ξ. It is not hard to see that $m_s \leqslant n_s$. We shall say that $W_1, W_2, \ldots, W_n$ are free up to step s at ξ if $m_s = n_s$. In particular, the condition for $s=1$ means simply that $W_1|_\xi, W_2|_\xi, \ldots, W_n|_\xi$ are linearly independent. Furthermore, note that if $W_1, W_2, \ldots, W_n$ are free up to step s at ξ, then they are free up to step s at η for all η in a neighborhood of ξ.

We shall now lift our original vector fields $\{X_k\}_{k=1,2,\ldots,n}$, locally, to a higher dimensional space in such a way that they become free up to step r. We shall also write $\tilde{m}$ for the dimension of the free nilpotent Lie algebra of step r on n generators.

THEOREM 4. *Let $X_1, X_2, \ldots, X_n$ be vector fields on a manifold M of dimension m such that the commutators of length $\leqslant r$ span the tangent space at $\xi \in M$. Then in terms of new variables, $t_{m+1}, t_{m+2}, \ldots, t_{\tilde{m}}$, there exist smooth functions $\lambda_{kl}(\eta, t)$ defined in a neighborhood $\tilde{U}$ of $\tilde{\xi} = (\xi, 0) \in M \times \mathbf{R}^{\tilde{m}-m} = \tilde{M}$ such that the vector fields $\{\tilde{X}_k\}$ given by*

$$\tilde{X}_k = X_k + \sum_{l=m+1}^{\tilde{m}} \lambda_{kl}(\eta, t) \frac{\partial}{\partial t_l} \tag{7.2}$$

are free up to step r at every point in $\tilde{U}$.

The proof of Theorem 4 will require a precise notion of approximating an arbitrary vector field on a graded nilpotent Lie group N by a vector field which is left-invariant. If $\mathfrak{N} = \sum_{j=1}^{s} V^j$ is the Lie algebra of N, a choice of basis $\{Y_{jk}\}$ of V^j for each j gives rise to a coordinate system

$$(u_{jk}) \leftrightarrow \exp\left(\sum u_{jk} Y_{jk}\right).$$

For a multi-index $\alpha = (j_1k_1, j_2k_2, \ldots, j_lk_l)$, $|\alpha| = \sum_{i=1}^{s} j_i$, we write u_α for $u_{j_1k_1} u_{j_2k_2} \ldots u_{j_lk_l}$ and D^α for $\partial^l/(\partial u_{j_1k_1} \partial u_{j_2k_2} \ldots \partial u_{j_lk_l})$. In terms of the definition of degree of homogeneity given in § 5, if α, β are multi-indices $u_\alpha D^\beta$ is a homogeneous differential operator of degree $|\beta| - |\alpha|$. For a general differential operator on N a notion of local degree (at 0) on N is defined as follows. If $f_\alpha(u)$ is smooth on N we shall say that the differential operator $f_\alpha(u) D^\alpha$ is of *(local) degree* $\leqslant \lambda$ if the Taylor expansion $f_\alpha(u) D^\alpha \sim \sum_\beta c_{\alpha\beta} u_\beta D^\alpha$ around $u=0$ is a formal sum of homogeneous differential operators of degree $\leqslant \lambda$. More generally, a smooth differential operator $D = \sum_\alpha f_\alpha(u) D^\alpha$ is of local degree $\leqslant \lambda$ if each $f_\alpha(u) D^\alpha$ is.

Now suppose $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_n$ are free up to step s at a point $\tilde{\xi} \in \tilde{M}$, and that the commutators of length $\leqslant s$ span the tangent space. For each j, $1 \leqslant j \leqslant s$, choose $\{\tilde{X}_{jk}\}_k$, commutators of length j with $\tilde{X}_{1k} = \tilde{X}_k$, $k = 1, 2, \ldots, n$, such that $\{\tilde{X}_{jk}\}_{j,k}$ restricted to $\tilde{\xi}$ is a basis of $T_{\tilde{\xi}}(\tilde{M})$. Then $\{\tilde{X}_{jk}\}_{j,k}$ determines a system of coordinates around $\tilde{\xi}$ by the exponential map based at $\tilde{\xi}$:

$$(u_{jk}) \leftrightarrow \exp\left(\sum u_{jk} \tilde{X}_{jk}\right) \cdot \tilde{\xi}. \tag{7.3}$$

This will be called a *system of canonical coordinates*; it is dependent on the choice of basis $\{\tilde{X}_{jk}\}_{j,k}$.

We may now give our main result comparing $\tilde{X}_k$ with Y_k. $\{Y_k\}_{k=1,2,\dots,n}$ generate the free Lie algebra of step r as described in example 4, § 3.

THEOREM 5. *Let $\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_n$ be vector fields on a manifold $\tilde{M}$, $\tilde{\xi}_0 \in \tilde{M}$ such that*

(i) *commutators of length $\leqslant r$ span the tangent space, and*

(ii) *$\{\tilde{X}_k\}$ is free up to step r at ξ_0.*

Choose $\{\tilde{X}_{jk}\}$, commutators of length $\leqslant r$, determining a system of canonical coordinates (u_{jk}) around $\tilde{\xi}_0$ by (7.3). Let $N = N_{F,r}$ be the free Lie group of step r on n generators and $\mathfrak{N}$ its Lie algebra. Then there is a basis $\{Y_{jk}\}$ of $\mathfrak{N}$ and neighborhoods $\tilde{V}$ of $\tilde{\xi}_0 \in \tilde{M}$ and U of $0 \in N$ with the following properties. On $\tilde{V} \times \tilde{V}$ consider the mapping to U

$$\Theta(\tilde{\xi}, \tilde{\eta}) = \exp\left(\sum u_{jk} Y_{jk}\right) \in U, \tag{7.4}$$

where $\tilde{\eta} = \exp\left(\sum u_{jk} \tilde{X}_{jk}\right)\tilde{\xi}$. Then for each fixed $\tilde{\xi}$ the mapping

$$\eta \to \Theta_{\tilde{\xi}}(\tilde{\eta}) = \Theta(\tilde{\xi}, \tilde{\eta}) = (u_{jk})$$

is a coordinate chart for V centered at $\tilde{\xi}$.

In this coordinate system

$$\tilde{X}_k = Y_k + R_k, \quad k = 1, \dots, n, \tag{7.5}$$

where R_k is a differential operator of local degree $\leqslant 0$.

Combining Theorems 4 and 5 we obtain the following local results for $\{X_k\}$ on a manifold M such that commutators of length $\leqslant r$ span at each point. Fixing $\xi_0 \in M$, Theorem 5 gives prolongations $\{\tilde{X}_k\}$ of $\{X_k\}$ with $\{\tilde{X}_k\}$ vector fields on $M \times \mathbf{R}^{\tilde{m}-m}$, $\tilde{m} = \dim N$ such that $\tilde{X}_k = X_k$ on functions constant in the new variables t_j. Furthermore, applying Theorem 5 shows that in local coordinates the $\tilde{X}_k$ thus obtained differ at each point $\tilde{\xi}$ from the generators of a free nilpotent Lie algebra by operators of local degree $\leqslant 0$.

§ 8. The main induction step.

We begin by giving a generalization of Theorem 5 which will be needed as well for the proof of Theorem 4. For what follows we shall first have to extend the notion of systems of canonical coordinates introduced in § 7.

Suppose $\{W_k\}$ is a set of vector fields on a manifold M', $\xi \in M'$, such that

(i) the commutators of length $\leqslant s_1$ span tangent space at ξ, and

(ii) $\{W_k\}$ is free up to step $s \leqslant s_1$.

A *partial system* of *canonical coordinates* around ξ is determined by a choice of a set of commutators $\{W_{jk}\}_{j,k}$, with $W_{1k}=W_k$. W_{jk} is a commutator of the $\{W_k\}$ of length j, such that $\{W_{jk}|_\xi\}_{j,k}$ is a basis of the tangent space and also satisfies the condition that $\{W_{jk}|_\xi\}_{j\leqslant s}$ spans the commutators of length $\leqslant s$ at ξ. This second condition may be satisfied by choosing each W_{jk} with j as small as possible. The coordinate system is then given by

$$(u_{jk})\leftrightarrow \exp(\sum u_{jk} W_{jk})\cdot \xi.$$

Now suppose $Y_{11}, Y_{12}, \ldots, Y_{1n}$ are generators of the free nilpotent Lie algebra $\mathfrak{N}_{F,s_1}$ of step s_1. (We recall that with respect to the dilations of $\mathfrak{N}_{F,s_1}$, described in § 5, each Y_{1k} is homogeneous of degree 1.) Then the correspondence

$$W_{1k}\leftrightarrow Y_{1k}=\varphi(W_{1k})$$

extends to a one-to-one correspondence

$$W_{jk}\leftrightarrow Y_{jk}=\varphi(W_{jk}),\quad j\leqslant s,$$

where $\{Y_{jk}\}_{j\leqslant s}$ spans the commutators of length $\leqslant s$ of the $\{Y_{1k}\}$, and the W_{jk} are as above. Furthermore, this correspondence can be chosen so that if $W_{jk}=[W_{j-1,k_1}, W_{1k_2}]$

$$\varphi(W_{jk})=[\varphi(W_{j-1,k_1}), \varphi(W_{1k_2})]. \tag{8.1}$$

We shall say that the correspondence φ is a *partial isomorphism up to step s*.

For our generalization of Theorem 5 we shall assume that we are given vector fields $W_1, W_2, \ldots, W_n$ which are free up to step s, but for which commutators of length $s+1$ are needed to span the tangent space of a fixed point ξ. We shall assign to $\{W_k\}$ a nilpotent Lie algebra $\mathfrak{N}_\xi$ of step $s+1$, free up to step s. In case $W_1, \ldots, W_n$ are actually free up to step $s+1$, $\mathfrak{N}_\xi$ will turn out to be the free nilpotent Lie algebra of step $s+1$.

To simplify our construction we consider first the following example. Suppose $n=3$, $s+1=2$, W_1 W_2, W_3 are linearly independent (free up to step 1) and W_1, W_2, W_3 and $\{[W_i, W_j]\}$ span at ξ. Suppose that the only linear relation is given by

$$c_1[W_1, W_2]|_\xi+c_2[W_2, W_3]|_\xi+c_3[W_1, W_3]|_\xi=\sum b_i W_i|_\xi.$$

Then $\mathfrak{N}_\xi$ will be the 2-step Lie algebra with generators Y_1, Y_2, Y_3, the only relation being $c_1[Y_1, Y_2]+c_2[Y_2, Y_3]+c_3[Y_1, Y_3]=0$. Now consider the general case where $W_1, \ldots, W_n$ are free up to step s, but span at step $s+1$. We construct $\mathfrak{N}_\xi$ of step $s+1$ in such a way that the linear relations among the commutators of length $s+1$ correspond to the linear relations of the commutators of the W_{jk} of length $s+1$ at ξ, modulo lower terms. Let $Y_1, \ldots, Y_n$ generate $\mathfrak{N}_{F,s+1}$, the free Lie algebra of step $s+1$, and extend the correspondence $W_k=W_{1k}\leftrightarrow Y_{1k}=Y_k$ to a partial isomorphism up to step s given by $W_{jk}\leftrightarrow Y_{jk}$. Now let

$\mathfrak{N}_\xi = \mathfrak{N}_{F,s+1}/\mathfrak{J}$, where $\mathfrak{J}$ is the ideal spanned by

$$\{\sum_{k_1,k_2} a_{k_1k_2}[Y_{sk_1}, Y_{1k_2}]\text{: there exist } c_{jk} \text{ such that } \sum_{k_1,k_2} a_{k_1k_2}[W_{sk_1}, W_{1k_2}]|_\xi = \sum_{\substack{j\leqslant s\\ k}} c_{jk} W_{jk}|_\xi\}.$$

We identify Y_{jk} with its image in $\mathfrak{N}_\xi$ for $j \leqslant s$, and extend the partial isomorphism

$$W_{jk} \leftrightarrow Y_{jk}, \quad j \leqslant s$$

by choosing a basis $\{Y_{s+1,k}\}$ for the commutators of length $s+1$ in $\mathfrak{N}_\xi$. Then if

$$Y_{s+1,k} = [Y_{sk_1}, Y_{1k_2}], \text{ we let } Y_{s+1,k} \leftrightarrow W_{s+1,k}, \text{ where } W_{s+1,k} = [W_{sk_1}, W_{1k_2}].$$

By construction, $\{W_{jk}\}_{j\leqslant s+1}$ restrict to a basis of the tangent space at ξ.

(8.2) LEMMA. *Let $\{W_k\}$ be free up to step s, and assume the commutators of length $\leqslant s+1$ span at ξ. Construct the nilpotent Lie algebra $\mathfrak{N}_\xi$ as above and the correspondence $W_{jk} \leftrightarrow Y_{jk}$, where $\{Y_k\}$ generate $\mathfrak{N}_\xi$. Write each Y_k and W_k in the common coordinate system*

$$(u_{jk}) \leftrightarrow \exp(\sum_{j,k} u_{jk} Y_{jk}) \leftrightarrow \exp(\sum_{j,k} u_{jk} W_{jk}) \cdot \xi.$$

Then

$$W_k = Y_k + R_k,$$

where R_k is of local degree $\leqslant 0$ on the graded Lie group N_ξ corresponding to $\mathfrak{N}_\xi$.

This lemma, and other technical results, will be proved in § 11. Theorem 5 is an immediate consequence of Lemma 8.2 with $s+1=r$ since $\mathfrak{N}_\xi = \mathfrak{N}_{F,s+1}$ in case $\{W_k\}$ is actually free up to step $s+1$.

We now show that we can add vector fields in new variables to the generators $Y_1, Y_2, \ldots, Y_n$ of $\mathfrak{N}_\xi$ to produce $\tilde{Y}_1, \ldots, \tilde{Y}_n$ free up to step $s+1$. More precisely, we have the following.

(8.3) LEMMA. *There exist smooth functions $\gamma_{kl}(x, t)$, $x \in N_\xi$, $t \in \mathbf{R}^q$ such that*

$$\tilde{Y}_k = Y_k + \sum_{l=1}^{q} \gamma_{kl}(x,t)\frac{\partial}{\partial t_l}, \quad k = 1, 2, \ldots, n,$$

are free up to step $s+1$ and are such that the commutators of length $\leqslant s+1$ span the tangent space of $N_\xi \times \mathbf{R}^q$.

Lemma 8.3 is proved by the following more general result, whose proof will be given in § 11.

(8.4) LEMMA. *Let $\mathfrak{G}$ be a Lie algebra, $\mathfrak{J}$ an ideal of $\mathfrak{G}$, $\mathfrak{G}_1$ the quotient $\mathfrak{G}_1=\mathfrak{G}/\mathfrak{J}$ and π: $\mathfrak{G}\to\mathfrak{G}_1$ the projection. Let $\tilde{Y}_1, \tilde{Y}_2, ..., \tilde{Y}_{k_1}\in\mathfrak{G}$ be chosen so that $\pi(\tilde{Y}_1), \pi(\tilde{Y}_2), ..., \pi(\tilde{Y}_{k_1})$ is a vector space basis of $\mathfrak{G}_1$ and let $Z_1, Z_2, ..., Z_{k_2}$ be a basis of $\mathfrak{J}$. Choose canonical coordinates in a neighborhood of the identities of the connected groups G, G_1 corresponding to $\mathfrak{G}$, $\mathfrak{G}_1$ respectively:*

$$(u, v) = ((u_i), (v_j)) \leftrightarrow \exp\left(\sum_{i=1}^{k_1} u_i \tilde{Y}_i + \sum_{j=1}^{k_2} v_j Z_j\right)$$

on G, and

$$u = (u_i) \leftrightarrow \exp\left(\sum_{i=1}^{k} u_i \pi(\tilde{Y}_i)\right).$$

Then there are smooth functions a_{ip} on $\mathbf{R}^{k_1}$ and b_{jq} on $\mathbf{R}^{k_2}$ such that

$$\tilde{Y}_j = \sum_p a_{jp}(u)\frac{\partial}{\partial u_p} + \sum_q b_{jq}(u, v)\frac{\partial}{\partial v_q}$$

and

$$\pi(\tilde{Y}_j) = \sum_p a_{jp}(u)\frac{\partial}{\partial u_p}$$

near 0.

Finally we combine Lemmas 8.2 and 8.3 to lift the $\{W_j\}$ to $\{\tilde{W}_j\}$ which are free up to step $s+1$. The result we prove is the following.

(8.5) LEMMA. *Let $\{W_k\}$ be free up to step s with commutators of length $\leqslant s+1$ spanning at ξ. Let $\{Y_k\}$ be the corresponding generators of the nilpotent Lie algebra $\mathfrak{N}_\xi$ as in Lemma 8.2. Lift $\{Y_k\}$ to new vector fields $\{\tilde{Y}_k\}$, $\tilde{Y}_k = Y_k + \sum_l \gamma_{kl}(x, t)\partial/\partial t_l$, $t\in\mathbf{R}^q$, by Lemma 8.3. Regarding each W_k as a vector field on N_ξ, the Lie group of $\mathfrak{N}_\xi$, put $\tilde{W}_k = W_k + \sum\gamma_{kl}(x, t)\partial/\partial t_l$. Then $\{\tilde{W}_k\}$ is free up to step $s+1$, and commutators of length $\leqslant s+1$ span the (higher dimensional) tangent space.*

Proof. Using the coordinate systems given in Lemmas 8.2 and 8.4 we may regard $\tilde{W}_k$ and $\tilde{Y}_k$ as vector fields on $N_{F,s+1}$. Then by Lemma 8.2, $\tilde{W}_k = \tilde{Y}_k + R_k$, where R_k is of degree $\leqslant 0$. (Note that if an operator on N_ξ is of degree $\leqslant 0$, then it is still of degree $\leqslant 0$ when regarded as an operator on $N_{F,s+1}$ by the correspondence of coordinate systems given in Lemma 8.4.)

Now let $\{\tilde{W}_{jk}\}$, $\{\tilde{Y}_{jk}\}$ be the corresponding commutators of length j of the $\tilde{W}_k$ and $\tilde{Y}_k$ respectively. We claim that

$$\tilde{W}_{jk} = \tilde{Y}_{jk} + R_{jk}, \tag{8.6}$$

where R_{jk} is of local degree $\leqslant j-1$, which will prove the lemma. Indeed, this is true for

$j=1$, where $W_{1k}=W_k$, $Y_{1k}=Y_k$. Now (8.6) follows easily by induction, using the simple fact that if D_1, D_2 are differential operators of local degrees less than or equal to λ_1 and λ_2 respectively, then $[D_1, D_2]$ is of local degree $\leqslant \lambda_1+\lambda_2$.

§ 9. Proof of Theorem 4.

Now assume $X_1, \ldots, X_n$ are real vector fields for which the commutators of length $\leqslant r$ span the tangent space at a fixed point ξ. We show first how to lift the $\{X_k\}$ to $\{X_k^{(1)}\}$ which are linearly independent (=free up to step 1) at ξ. Then we shall use the results of § 8 show that vector fields free up to step s can be lifted to be free up to step $s+1$.

Rearranging the X_k's if necessary we may assume $X_1, X_2, \ldots, X_p$, $p\leqslant n$, is a maximal set of linearly independent vectors among the X_k at ξ, $k=1, \ldots, n$. Put

$$X_{1k}^1 = X_k, \quad k\leqslant p \tag{9.1}$$
$$X_{1,p+l}^1 = X_{p+l} + \frac{\partial}{\partial t_{m+l}}, \quad l=1,2,\ldots,n-p.$$

The commutators of length $s\geqslant 2$ of the $\{X_{1k}^1\}$ are the same as that of the $\{X_k\}$. Furthermore, $X_{1k}^1|_{\xi_1}$, $k=1, \ldots, n$, are linearly independent, where $\xi_1=(\xi, 0)\in U\times \mathbf{R}^{m-p}$, and by dimensional considerations their commutators of order $\leqslant r$ span $T_{\xi_1}(U\times\mathbf{R}^{m-p})$.

Assume now by induction that for some $s\geqslant 1$ one can find vector fields $F_k^s = \sum_{p=m+1}^{q}\lambda_{kp}^s(\eta, t)\partial/\partial t_p$, where $t\in\mathbf{R}^{q-m}$ and $\lambda_{kp}^s(\eta, t)$ are defined and smooth for all (η, t) in a neighborhood of $(\xi, 0)\in M\times\mathbf{R}^{q-m}$ such that

(9.2) (i) $X_k^s = X_k + F_k^s$, $k=1, \ldots, n$, are free up to step s, and
(ii) the commutators $\{X_{jk}^s\}$, of length $\leqslant r$ span the tangent space at $\xi_s=(\xi, 0)$.

If $s+1<r$, Lemma 8.4 cannot be applied directly to lift the $\{X_k^s\}$ to $\{X_k^{s+1}\}$, to be free up to step $s+1$. Therefore, in this case we shall restrict the X_k^1 to act on a lower dimensional space. To do this we shall need specific information about the X_k^s in a partial canonical coordinate system. Let $u=(u_{jk})\leftrightarrow\exp(\sum u_{jk}X_{jk}^s)\cdot\xi$ be such a system. In analogy with the case where u is on a graded group we define the family δ_t of dilations by $\delta_t(u_{jk})=(t^j u_{jk})$, and we define homogeneous functions as before with respect to δ_t. Note, however, that unlike the group case degrees of homogeneity here may depend on the choice of partial canonical coordinates if $s+1<r$. We shall write $O(|u|^l)$ for any function whose Taylor series at $u=0$ is a formal sum of homogeneous terms of degree $\geqslant l$.

(9.3) Lemma. *Suppose X_{1k}^s, $k=1, 2, \ldots, n$, are free up to step s, where commutators of length $\leqslant r$ span the tangent space at ξ_s, Let $\{X_{jk}^s\}$ be commutators determining a partial canonical* *coordinate system around* ξ_s,

$$u=(u_{jk})\leftrightarrow \exp(\sum_{j,k} u_{jk}X^s_{jk})\cdot \xi_s.$$

Then if k' is fixed, $1\leqslant k'\leqslant n$,

$$X^s_{1k'}=\sum_{j\leqslant s}(g^{k'}_{jk}(u)+e^{k'}_{jk}(u))\frac{\partial}{\partial u_{jk}}+\sum_k g^{k'}_{s+1,k}(u)\frac{\partial}{\partial u_{s+1,k}}+\sum_{j,k}O(|u|^{s+1})\frac{\partial}{\partial u_{jk}},$$

where $g^{k'}_{jk}$ is homogeneous of degree $j-1$, and $e^{k'}_{jk}$ is homogeneous of degree s.

This lemma will be proved in § 11. Using it, we define the vector fields $W^s_{1k'}$, $k'=1,2,\ldots,n$, by

$$W^s_{1k'}=\sum_{j\leqslant s}(g^{k'}_{jk}(u)+e^{k'}_{jk}(u))\frac{\partial}{\partial u_{jk}}+\sum_k g^{k'}_{j+1,k}(u)\frac{\partial}{\partial u_{s+1,k}}.$$

We think of the W^s_{1k} as acting on the Euclidean space with coordinates (u_{jk}), $j\leqslant s+1$. Now we claim that $\{W^s_{1k}\}$ is free up to step s, and that the commutators up to step $s+1$ span the tangent space. To prove this, let W^s_{jk}, $1\leqslant j\leqslant s+1$, be the commutators of the W^s_{1k} corresponding to the X^s_{jk}, $j\leqslant s+1$. We shall prove

$$W^s_{j_0k_0}|_{u=0}=X^s_{j_0k_0}|_{u=0}, \tag{9.4}$$

for all $j_0\leqslant s+1$. Indeed, applying (9.4) with $j_0\leqslant s$ will show that $\{W^s_{1k}\}$ is free up to step s, since $\{X^s_{1k}\}$ is free up to step s. Next, since $X^s_{j_0k_0}|_{u=0}=\partial/\partial u_{j_0k_0}$, (9.4) applied to all $j_0\leqslant s+1$ show that $\{W^s_{j_0k_0}|_{u=0}\}$ spans the tangent space to the coordinates u_{jk}, $j\leqslant s+1$.

To see (9.4) note first that by definition $W^s_{1k}=X^s_{1k}-\sum_{j,k}O(|u|^{s+1})\partial/\partial u_{k}$. We claim that for any $j_0\leqslant s+1$,

$$W^s_{j_0k_0}=X^s_{j_0k_0}-\sum_{j,k}O(|u|^1)\frac{\partial}{\partial u_{jk}}. \tag{9.5}$$

To prove this, note that

$$\left[O(|u|^{s+1})\frac{\partial}{\partial u_{j_1k_1}},\, g^{k'}_{j_2k_2}(u)\frac{\partial}{\partial u_{j_2k_2}}\right]$$
$$=-g^{k'}_{j_2k_2}(u)\frac{\partial}{\partial u_{j_2k_2}}(O(|u|^{s+1}))\frac{\partial}{\partial u_{j_1k_1}}+O(|u|^{s+1})\frac{\partial}{\partial u_{j_1k_1}}(g^{k'}_{j_2k_2}(u))\frac{\partial}{\partial u_{j_2k_2}}.$$

The first term on the right has a coefficient which is $O(|u|^{(j_2-1)+(s+1-j_2)})=O(|u|^s)$, while the second term has a coefficient which is at least $O(|u|^{s+1})$. Similarly

$$\left[O(|u|^{s+1})\frac{\partial}{\partial u_{j_1k_1}},\, e^{k'}_{j_2k_2}\frac{\partial}{\partial u_{j_2k_2}}\right]$$

has coefficients which are $O(|u|^s)$. Applying this argument s times one obtains (9.5), and hence (9.4).

We may now apply Lemma 8.5 to the vector fields W^s_{1k}. Thus we can find vector fields $F^s_k = \sum_l \gamma^s_{kl}(u, t)\partial/\partial t_l$, $t \in \mathbf{R}^{q_l}$, such that by defining $W^{s+1}_{1k} = W^s_{1k} + F^s_k$, we have that W^{s+1}_{1k}, $k = 1, \dots, n$, are free up to step $s+1$, and that their commutators of lengths $\leqslant s+1$ span the tangent space.

Now put

$$X^{s+1}_k = X^s_k + F^s_k, \quad k = 1, 2, \dots, n.$$

In order to complete the proof of Theorem 4 it will suffice to show that $\{X^{s+1}_k\}$ is free up to step $s+1$, and that the commutators of lengths $\leqslant r$ span the tangent space at $(\xi_s, 0)$ in the extended space. Let $\{X^{s+1}_{jk}\}$ be the commutators of the X^{s+1}_k corresponding to the $\{X^s_{jk}\}$, and $\{W^{s+1}_{jk}\}$ that of the $W^{s+1}_{1k} = W^{s+1}_k$. To prove that $\{X^{s+1}_k\}$ is free up to step $s+1$ it will suffice to show that

$$X^{s+1}_{jk}|_{u=0} = W^{s+1}_{jk}|_{u=0} \tag{9.6}$$

for all $j \leqslant s+1$. The proof of this is completely analogous to that of (9.4) because

$$\left[O(|u|^{s+1})\frac{\partial}{\partial u_{j_1 k_1}},\ \gamma^s_{kl}(u, t)\frac{\partial}{\partial t_l}\right] = O(|u|^{s+1})\frac{\partial}{\partial u_{j_1 k_1}}(\gamma^s_{kl})\frac{\partial}{\partial t_l},$$

since $O(|u|^{s+1})$ does not involve the variables t_l.

Now we shall show that the X^{s+1}_{jk} span. First we claim that for any j,

$$X^{s+1}_{jk} = X^s_{jk} + F_{jk}, \tag{9.7}$$

where F_{jk} involves differentiation only in the new variables. To see (9.7) it suffices, since $X^{s+1}_{1k} = X^s_{1k} + F^s_k$, to note that if $h(u)$ does not vary with the new t_l's, then

$$\left[h(u)\frac{\partial}{\partial u_{jk}},\ \gamma^s_{kl}(u, t)\frac{\partial}{\partial t_l}\right] = h(u)\frac{\partial}{\partial u_{jk}}\gamma^s_{kl}(u, t))\frac{\partial}{\partial t_l}.$$

Now (9.6) shows

$$\dim(\operatorname{span} X^{s+1}_{jk}|_{u=t=0}, j \leqslant s+1) = \dim(\operatorname{span} X^s_{j,k}|_{u=0}, j \leqslant s+1) + \text{number of added variables.}$$

Hence, because of (9.7)

$$\dim(\operatorname{span} X^{s+1}_{jk}|_{u=t=0}) \geqslant \dim(\operatorname{span} X^s_{jk}|_{u=0}) + \text{number of added variables.}$$

Since the reverse inequality is obvious, we have proved the spanning property.

§ 10. The Campbell-Hausdorff formula

We shall prove Lemmas 8.2, 8.4, and 9.3 by explicit calculations of vector fields in canonical coordinates. Before giving the proofs, we need to make some preliminary remarks on the application of the Campbell-Hausdorff formula.

Let $\xi \in U$, where U is an open subset of $\mathbf{R}^l$, and suppose $\{W_{jk}\}_{j,k}$ are real vector fields which form a basis of T_ξ, the tangent space at ξ. With U replaced by a smaller neighborhood of ξ if necessary, the exponential map exp identifies each $u=(u_{jk})$ with $\sum_{j,k}|u_{jk}|^2$ sufficiently small, with a transformation $\exp u\cdot W$, where $u\cdot W=\sum_{j,k}u_{jk}W_{jk}$. The image of an element $\eta\in U$ under this transformation will be denoted $\exp(u\cdot W)\eta$. Now if ξ is fixed and f is a smooth function defined in a neighborhood of ξ, then f may be regarded as a function of u near ξ; for any η near ξ, $\eta=\exp(u\cdot W)\xi$, uniquely for some $u=(u_{jk})$. It might be useful in this section to distinguish between f as a function of η and f as a function of u. Thus we define f' near $u=0$ by

$$f'(u)=f((\exp u\cdot W)\cdot\xi).$$

We claim that the Taylor expansion of f' around $u=0$ is given by the formal power series $(e^{(u\cdot W)}f)(\xi)$. Indeed, this follows from the identity

$$\frac{\partial}{\partial u_{jk}}(f'(u))=W_{jk}f(\exp u\cdot W)\cdot\xi). \tag{10.1}$$

The correspondence $f\leftrightarrow f'$ on functions gives rise to a correspondence $X\leftrightarrow X'$ of vector fields, defined by $X'f'(u)=Xf(\exp(u\cdot W)\xi)$. We want to calculate the Taylor series of X' in local coordinates around $u=(0)$. By the above, for any f', the Taylor series of $X'f'$ around $u=(0)$ is given by

$$X'f'(u)\sim e^{u\cdot W}Xf(\xi), \tag{10.2}$$

where the right hand side is interpreted as a formal power series in u. Now fix (j_0, k_0) and let $f=h_{j_0k_0}$ be the coordinate function $h_{j_0k_0}(\eta)=u_{j_0k_0}$ if $\eta=\exp(u\cdot W)\xi$. Then $e^{u\cdot W}Xh_{j_0k_0}(\xi)$ is the Taylor expansion of the coefficient of $\partial/\partial u_{j_0k_0}$ at $u=0$. That is, if $X'=\sum a'_{jk}(u)\partial/\partial u_{jk}$, then

$$a'_{j_0k_0}(u)\sim e^{u\cdot W}Xh_{j_0k_0}(\xi). \tag{10.3}$$

In order to calculate the right hand side of (10.3) we consider the formal power series in u and τ given by

$$e^{u\cdot W}e^{\tau X}f(\xi). \tag{10.4}$$

We may consider (10.4) as a formal power series in τ whose coefficients are formal power series in u. The right hand side of (10.2) is then obtained as the coefficient of τ in (10.4). We express this equality of power series symbolically by

$$e^{u\cdot W}Xf(\xi)\sim\frac{d}{d\tau}(e^{u\cdot W}e^{\tau X}f(\xi))|_{\tau=0}. \tag{10.5}$$

In order to make use of the formula (10.5) we shall have to express $e^{u\cdot W}e^{\tau X}$ as an exponential. This is done formally by the Campbell-Hausdorff formula: If x and y are non-commuting indeterminates, the following is an equality of formal power series in x and y, e.g., [21], LA 4.17:

$$e^{x}e^{y}=e^{x+y+h(x,y)}. \tag{10.6}$$

Here $h(x, y)$ is a formal sum over ordered pairs of positive integers (j, k) of terms

$$h_{(j,k)}(x,y)=\sum_{\substack{p_i+q_i>0\\ 2\leqslant\Sigma p_i+q_i\leqslant k}} c(p,q)\,D_{p_1,\dots,p_j,q_1,\dots,q_j}(x,y)$$

where $c(p, q)$ is a constant,

$$D_{p_1,\dots,p_j,q_1,\dots,q_j}(x,y)=\begin{cases}(\mathrm{Ad}\,x)^{p_1}(\mathrm{Ad}\,y)^{q_1},\ \dots,\ (\mathrm{Ad}\,x)^{p_j}(\mathrm{Ad}\,y)^{q_j-1}(x) & \text{if } q_j\geqslant 1\\ (\mathrm{Ad}\,x)^{p_1}(\mathrm{Ad}\,y)^{q_1},\ \dots,\ (\mathrm{Ad}\,y)^{q_j-1}(\mathrm{Ad}\,x)^{p_j-1}(y) & \text{if } q_j=0\end{cases}$$

with $(Ada)(b)$ defined as $[a\ b]$.

We consider the application of this formula to calculate the right hand side of (10.5) for $X=W_{1k_1}$, $1\leqslant k\leqslant n$, $f=h_{j_0k_0}$. Suppose that we are only interested in the Taylor expansion up to an error which is $O(|u|^{l})$, in terms of the homogeneity defined in § 9. By (10.6)

$$e^{u\cdot W}e^{\tau W_{1k_1}}=\exp(u\cdot W+\tau W_{1k_1}+\tau(\sum_{1\leqslant p<l}c_p(\mathrm{Ad}\,u\cdot W)^p\cdot W_{1k_1})+O(|u|^l,\tau^2)).$$

Thus,

$$e^{u\cdot W}e^{\tau W_{1k_1}}h_{j_0k_0}(\xi)=\exp(u\cdot W+\tau W_{1k_1}+\tau(\sum_{1\leqslant p<l}c_p(\mathrm{Ad}\,u\cdot W)^pW_{1k_1}))\,h_{j_0k_0}(\xi)+O(|u|^l,\tau^2). \tag{10.7}$$

Note also that since $u\cdot W+\tau W_{1k_1}+\tau(\sum_{p<l}c_pAd(u\cdot W)^pW_{1k_1})$ is an actual vector field (rather than merely a formal power series), the coefficient of τ in the first term on the right of (10.7) is the actual derivative

$$\frac{d}{d\tau}(h_{j_0k_0}(\exp(u\cdot W+\tau W_{1k_1}+\tau(\sum_{1\leqslant p<l}c_p(\mathrm{Ad}\,u\cdot W)^p\cdot W_{1k_1}))\cdot\xi))|_{\tau=0}.$$

§ 11. Proofs of the Lemmas of § 8 and § 9.

Our main technique in these proofs will be the use of the Campbell-Hausdorff formula as developed in § 10.

Proof of Lemma 8.2. We assume $W_{1k}=W_k$, $k=1,2,\dots,n$, are free up to step s at ξ and that commutators of length $\leqslant s+1$ span. Let $W_k\leftrightarrow Y_{jk}$, $\{Y_{jk}\}$ spanning $\mathfrak{N}_\xi$, determine the common coordinate system

$$u=(u_{jk})\leftrightarrow\exp(\sum_{j,k}u_{jk}Y_{jk})\leftrightarrow\exp(\sum_{j,k}u_{jk}W_{jk}).$$

In order to prove $W_{k'} = Y_{k'} + R_{k'}$, it suffices to show that if

$$Y_{k'} = \sum_{j,k} f^*_{jk}(u) \frac{\partial}{\partial u_{jk}},$$

and

$$W_{k'} = \sum_{j,k} f_{jk}(u) \frac{\partial}{\partial u_{jk}},$$

then

$$f_{jk}(u) = f^*_{jk}(u) + O(|u|^j). \tag{11.1}$$

Here, as in § 9, $O(|u|^j)$ indicates a function whose Taylor series at 0 is a sum of terms homogeneous of degrees $\geqslant j$. Fix $(j_0 k_0)$ and let $h_{(j_0k_0)}$, $h^*_{j_0k_0}$ be the coordinate functions around ξ and 0 respectively:

$$h_{j_0k_0}((\exp u \cdot W)\,\xi) = u_{j_0k_0},$$

$$h^*_{j_0k_0}(\exp u \cdot Y) = u_{j_0k_0},$$

where $u \cdot W = \sum_{j,k} u_{jk} W_{jk}$, $u \cdot Y = \sum_{j,k} u_{jk} Y_{jk}$. By (10.3), (10.5)

$$f_{j_0k_0}(u) \sim \frac{d}{d\tau}(e^{u\cdot W} e^{\tau W_{1k'}} h_{j_0k_0}(\xi))|_{\tau=0}. \tag{11.2}$$

By (10.7),

$$e^{u\cdot W} e^{\tau W_{1k'}} h_{j_0k_0}(\xi) = \exp(u \cdot W + \tau W_{1k'} + \tau(\textstyle\sum c_p(\mathrm{Ad}(u \cdot W))^p)\, W_{1k_1})\, h_{j_0k_0}(\xi) + O(|u|^{s+1}, \tau^2). \tag{11.3}$$

A similar calculation holds for $f^*_{j_0k_0}$, replacing W by Y and $h_{j_0k_0}$ by $h^*_{j_0k_0}$. Thus we must compare the expansions $\mathrm{Ad}(u\cdot W)^p W_{1k'}$ with $\mathrm{Ad}(u\cdot Y)^p Y_{1k'}$. If Q_{jk} denotes either W_{jk} or Y_{jk}, then

$$\mathrm{Ad}(u \cdot Q)^p \cdot Q_{1k_1} = \sum_\alpha b_{p\alpha} u_\alpha Q_\alpha, \quad \alpha = (j_1 k_1, j_2 k_2, \ldots, j_p k_p), \tag{11.4}$$

where $u_\alpha = u_{j_1k_1} \cdot u_{j_2k_2} \cdot \ldots \cdot u_{j_pk_p}$, and $Q_\alpha = [Q_{j_1k_1}, [Q_{j_2k_2}, [\,, \ldots, [Q_{j_pk_p}, Q_{1k'}], \ldots,]$.

Note that u_α is homogeneous of degree $|\alpha| = \sum_{i=1}^p j_i$, and Q_α is a commutator (of the Q_{1k}) of length $|\alpha| + 1$. In calculating (11.3) we compare the expansions of W_α and Y_α. It suffices to do this for $|\alpha| \leqslant s$, since $u_\alpha W_\alpha$, $|\alpha| \geqslant s+1$ will contribute a term which is $O(|u|^{s+1})$ which may therefore by absorbed in the error term.

If $|\alpha| \leqslant s-1$, then Q_α is a commutator of length $\leqslant s$. Hence by the partial isomorphism up to step s

$$Y_\alpha = \sum_k a_{\alpha k} Y_{|\alpha|+1,k}$$

$$W_\alpha = \sum_k a_{\alpha k} W_{|\alpha|+1,k}, \tag{11.5}$$

where the $a_{\alpha k}$ are constants. Now suppose $|\alpha| = s$. By the construction of $\mathfrak{N}_\xi$, if

$$Y_\alpha = \sum_k a_{\alpha k} Y_{s+1,k}, \tag{11.6}$$

then

$$W_\alpha = \sum_k a_{\alpha k} W_{s+1,k} + \sum_{\substack{j \leq s \\ k}} a'_{\alpha jk} W_{jk} + \sum_{j,k} z^\alpha_{jk}(\eta) W_{jk}, \tag{11.7}$$

where $z^\alpha_{jk}(\xi) = 0$. Substituting this into (11.3) gives

$$\begin{aligned} e^{u \cdot W} e^{\tau W_{1k'}} h_{j_0 k_0}(\xi) = \exp(u \cdot W + \tau W_{1k'} + \tau(\sum_{j \leq s} (g_{jk}(u) + e_{jk}(u)) W_{jk} \\ + \sum_k g_{s+1,k}(u) W_{s+1,k}) + \tau O(|u|^s) \sum z^\alpha_{jk} W_{jk}) h_{j_0 k_0}(\xi) + O(|u|^{s+1}, \tau^2), \end{aligned} \tag{11.8}$$

where g_{jk} is homogeneous of degree $j-1$ if $j \geq 2$, $g_{1k} = 0$, e_{jk} is homogeneous of degree s, and the $z^\alpha_{jk}(\eta)$ are smooth functions such that $z^\alpha_{jk}(\xi) = 0$. (Note that the e_{jk} come from the second term on the right in (11.7), multiplied by u_α.)

Now we claim that the contribution of the term $\tau O(|u|^s) \sum_{j,k|\alpha|=s} z^\alpha_{jk}(\eta) W_{jk}$, appearing in the exponential of (11.8), may be absorbed in $O(|u|^{s+1}, \tau^2)$. Indeed, since

$$(z^\alpha_{jk}(\eta) W_{jk}) h_{j_0 k_0}(\xi) = z^\alpha_{jk}(\xi)(W_{jk} h_{j_0 k_0})(\xi) = 0,$$

this term contributes only a product with other terms in the exponential, when the right hand side of (11.8) is expanded. Since any product of $\tau O(|u|^s) z^\alpha_{jk}(\eta) W_{jk}$ with one of the other terms is $O(|u|^{s+1}, \tau^2)$, the claim is proved. Thus

$$\begin{aligned} &(\exp u \cdot W)(\exp \tau W_{1k'}) h_{j_0 k_0}(\xi) \\ &= \exp(u \cdot W + \tau W_{1k'} + \tau(\sum_{1 \leq j \leq s} (g_{jk} + e_{jk}) W_{jk} + \sum_k g_{s+1,k} W_{s+1,k})) h_{j_0 k_0}(\xi) + O(|u|^{s+1}, \tau^2). \end{aligned} \tag{11.9}$$

Now we claim that then

$$e^{p \cdot W} h_{j_0 k_0}(\xi) = p_{j_0 k_0}$$

for p sufficiently small. Indeed, by (10.2) the left side is the Taylor series, in the variables p_{jk}, of $h_{j_0 k_0}((\exp \sum_{j,k} p_{jk} W_{jk})\xi)$. By definition of $h_{j_0 k_0}$, this latter expression is equal to $p_{j_0 k_0}$, proving the claim. Applying this to the exponential term on the right in (11.9), we obtain

$$e^{u \cdot W} e^{\tau W_{1k'}} h_{j_0 k_0}(\xi) = \begin{cases} u_{j_0 k_0} + \tau + \tau(g_{j_0 k_0}(u) + e_{j_0 k_0}(u)) + O(|u|^{s+1}, \tau^2), \text{ if } (j_0 k_0) = (1, k') \\ u_{j_0 k_0} + \tau(g_{j_0 k_0}(u) + e_{j_0 k_0}(u)) + O(|u|^{s+1}, \tau^2), \text{ otherwise,} \end{cases}$$

where $e_{j_0 k_0}$ is defined to be zero for $j_0 \geq s+1$. Thus we obtain by (11.2)

$$f_{j_0 k_0}(u) = \begin{cases} 1 + g_{j_0 k_0}(u) + e_{j_0 k_0}(u) + O(|u|^{s+1}), \text{ if } (j_0, k_0) = (1, k') \\ g_{j_0 k_0}(u) + e_{j_0 k_0}(u) + O(|u|^{s+1}), \text{ otherwise.} \end{cases} \tag{11.10}$$

By a similar calculation for $Y_{1k'}$, using (11.5) and (11.6), we obtain

$$f^{*}_{j_0k_0}(u) = \begin{cases} 1 + g_{j_0k_0}(u) + O(|u|^{s+1}), & \text{if } (j_0, k_0) = (1, k') \\ g_{j_0k_0}(u) + O(|u|^{s+1}), & \text{otherwise.} \end{cases}$$

This completes the proof of Lemma 8.2.

Proof of Lemma 9.3. We again let $h_{j_0k_0}$ denote a coordinate function and follow the method of the proof of Lemma 8.2 to calculate $W_{k'}$ in partial canonical coordinates. It is easy to check that we still obtain the expression (11.10) for $f_{j_0k_0}(u)$, and by the homogeneities given there this suffices to prove Lemma 9.3.

Proof of Lemma 8.4. We compute $\tilde{Y}_k$ and $\pi(\tilde{Y}_k)$ in canonical coordinates by expanding

$$\exp\left(\sum_{i=1}^{k_1} u_i \tilde{Y}_i + \sum_{j=1}^{k_2} v_j Z_j\right) \exp(t\tilde{Y}_k) \tag{11.11}$$

and

$$\exp\left(\sum_{i=1}^{k_1} u_i \pi(\tilde{Y}_i)\right) \exp(t\pi(\tilde{Y}_k)), \tag{11.12}$$

using the Campbell-Hausdorff formula (10.6). Since $\mathcal{J}$ is an ideal, any commutator which contains a Z_{j_0} can be expressed as a linear combination of the Z'_j. Now the result can be obtained by comparing the expansions of (11.11) and (11.12). We omit the details.

This supplies the arguments omitted from § 8 and § 9, and thus completes the proofs of Theorems 4 and 5.

§ 12. Properties of the Map Θ.

For any point $\tilde{\xi}_0 \in \tilde{M} = M \times \mathbf{R}^{\tilde{m}-m}$, Theorem 5 produces a map Θ: $\tilde{V} \times \tilde{V} \to U$, where $\tilde{V}$ is a neighborhood of $\tilde{\xi}_0$ and U is a neighborhood of 0 in the free Lie group $N = N_{F,r}$ of step r on n generators. Recall that for $\eta \in \tilde{V}$ we define $\Theta\tilde{\eta}$: $\tilde{V} \to U$ by $\Theta\tilde{\eta}(\tilde{\xi}) = \Theta(\tilde{\xi}, \tilde{\eta})$. To simplify notation, we shall now assume that $\tilde{X}_k = X_k$, $k = 1, \ldots, n$, and omit the tildas.

Since Θ is defined by means of the exponential map we have

$$\Theta(\xi, \eta) = -\Theta(\eta, \xi) = \Theta(\xi, \eta)^{-1}. \tag{12.1}$$

Using (12.1) we may define a pseudo-metric ϱ on V by putting

$$\varrho(\xi, \eta) = |\Theta(\xi, \eta)|, \tag{12.2}$$

where $|\ |$ is the norm function on N given by (5.2). We now show that ϱ satisfies inequalities, similar to the triangle inequality, which will be useful later.

(12.3) PROPOSITION. *If $\xi, \eta, \zeta \in \tilde{V}$ with $\varrho(\xi, \eta)$ and $\varrho(\xi, \zeta)$ both $\leqslant 1$, we have*

(i) $|\Theta(\xi,\eta)-\Theta(\zeta,\eta)| \leqslant C_1(\varrho(\xi,\zeta)+\varrho(\xi,\zeta)^{1/r}\varrho(\xi,\eta)^{1-(1/r)})$,

(ii) $\varrho(\zeta,\eta) \leqslant C_2(\varrho(\xi,\zeta)+\varrho(\eta,\xi))$.

Note that this proposition is the analogue for the general case of Theorem 14.10 (d) of Folland-Stein [8]. However, the proof of (12.3) seems to be much more complicated than that of the corresponding result in [8].

Proof. We first show that (ii) follows from (i). By the inequality (5.3) on groups we have

$$\varrho(\zeta,\eta) = |\Theta(\zeta,\eta)| = |(-\Theta(\xi,\eta)+\Theta(\zeta,\eta))+\Theta(\xi,\eta)| \leqslant \gamma(|\Theta(\xi,\eta)| + |\Theta(\xi,\eta)-\Theta(\zeta,\eta)|)$$

for some $\gamma > 0$. Since

$$\varrho(\xi,\zeta)+\varrho(\xi,\zeta)^{1/r}\varrho(\xi,\eta)^{1-(1/r)} \leqslant 2(\varrho(\xi,\zeta)+\varrho(\xi,\eta)),$$

(ii) will be proved when (i) is.

To establish (i) we first prove

(i)′ $$|\Theta(\xi,\eta)-\Theta(\zeta,\eta)| \leqslant C_2'(\varrho(\xi,\zeta)+\sum_{1<j\leqslant r}\sum_{1\leqslant l<j}\varrho(\xi,\zeta)^{l/j}\varrho(\xi,\eta)^{(j-l)/j}).$$

For this we fix ζ and establish local coordinates around ζ. Let h_{jk} be the corresponding coordinate functions:

$$h_{jk}(\eta) = (\Theta\zeta(\eta))_{jk}.$$

Now we define $v=(v_{jk})$, $w=(w_{jk})$ as functions of ξ, η by

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

and

$$w = \Theta\xi(\zeta).$$

Since

$$(\exp v\cdot X)\xi = \eta,\quad (\exp w\cdot X)\xi = \zeta \quad (\text{where } v\cdot X = \textstyle\sum_{j,k} v_{jk}X_{jk},\ w\cdot X = \sum_{j,k} w_{jk}X_{jk}),$$

$$(\exp v\cdot X)(\exp(-w\cdot X))\zeta = \eta.$$

Hence

$$h_{jk}(\eta) = h_{jk}(\exp v\cdot X \exp(-w\cdot X)\zeta).$$

As in § 10 it follows that the Taylor series of h_{jk} as a function of v, w, around $v=0$, $w=0$ is given by

(12.4) $$e^{v\cdot X}e^{(-w\cdot X)}h_{jk}(\zeta),$$

where $e^{v\cdot X}$ and $e^{(-w\cdot X)}$ are regarded as formal power series in v and w respectively.

We now expand (12.4) using the Campbell-Hausdorff formula (10.6). Then

$$e^{v\cdot X}e^{-w\cdot X}=\exp((v-w)\cdot X+\sum_{|\alpha|<r}d_\alpha q_\alpha X^\alpha+O(\sum_{0<l<r}|v|^l|w|^{r-l})) \tag{12.5}$$

where $\alpha=(j_1k_1, j_2k_2, \dots, j_sk_s)$, $|\alpha|=\sum_{i=1}^s j_i$, $q_\alpha=q_{j_1k_1}q_{j_2k_2}\dots q_{j_sk_s}$, with each $q_{j_ik_i}=v_{j_ik_i}$ or $w_{j_ik_i}$, and $X^\alpha=[X_{j_1k_1}, [X_{j_2k_2}, [, \dots, [X_{j_{s-1}k_{s-1}}, X_{j_sk_s}], \dots]$. X^α is a commutator of length $|\alpha|$. To check that (12.5) holds, it suffices to show that if $|\alpha|\geqslant r$, then q_α may be absorbed in the error term $O(\sum_{0<l<r}|v|^l|w|^{r-l})$. By the definition of the norm on the group,

$$\|v_{jk}\|\leqslant|v|^j \quad\text{and}\quad \|w_{jk}\|\leqslant|w|^j, \tag{12.6}$$

where $\|\ \|$ indicates Euclidean norm. (See (5.2).) We shall show, more generally, that for any α,

$$\|q_\alpha\|=O(\sum_{0<l<|\alpha|}|v|^l|w|^{|\alpha|-l}). \tag{12.7}$$

By definition, $q_\alpha=(v_{l_1m_1}v_{l_2m_2}, \dots, v_{l_{t_1}m_{t_1}})(w_{l_1'm_1'}, \dots, w_{l_{t_2}'m_{t_2}'})$ with $\sum l_i+\sum l_i'=|\alpha|$. Observe, however, that since q_α is a coefficient of a commutator, it cannot be made up entirely of either v's or w's. Thus $0<\sum l_i<|\alpha|$. Putting $l=\sum l_i$ and applying (12.6) we get (12.7). From (12.5) and the above we have

$$\begin{aligned}e^{v\cdot X}e^{-w\cdot X}h_{j_0k_0}(\zeta)&=\exp((v-w)\cdot X+\sum_{|\alpha|<r}O(\sum_{0<l<|\alpha|}|v|^l|w|^{|\alpha|-l}))h_{j_0k_0}(\xi)+O(\sum|v|^r|w|^{r-l})\\&=v_{j_0k_0}-w_{j_0k_0}+\sum_{0<l<j_0}O(|v|^l|w|^{j_0-l})+O(\sum_{0<l<r}|v|^l|w|^{r-l}).\end{aligned}$$

Therefore,

$$\begin{aligned}(\Theta(\xi,\eta)-\Theta(\zeta,\eta))_{j_0k_0}&=v_{j_0k_0}-(v_{j_0k_0}-w_{j_0k_0}+\sum_{0<l<j_0}O(|v|^l|w|^{j_0-l}))\\&=w_{j_0k_0}+\sum_{0<l<j_0}O(|v|^l|w|^{j_0-l}).\end{aligned} \tag{12.8}$$

Now since $|(u_{jk})|\leqslant C\sum_{1\leqslant j\leqslant r}\sum_k\|u_{jk}\|^{(1/j)}$, (i)′ follows from (12.8). To prove (i), it suffices to show that there exists a constant C' such that if A and B are positive real numbers,

$$A^{a_1}B^{1-a_1}\leqslant C'(A+A^aB^{1-a}) \tag{12.9}$$

whenever $0<a\leqslant a_1<1$. Indeed, if (12.9) holds we put $A=\varrho(\eta,\zeta)$, $B=\varrho(\xi,\eta)$, $a=1/r$ to obtain

$$\varrho(\eta,\zeta)^{1/j}\varrho(\xi,\eta)^{(j-1)/j}\leqslant C(\varrho(\eta,\zeta)+\varrho(\eta,\zeta)^{1/r}\varrho(\xi,\eta)^{(r-1)/r})$$

from which (i) will follow, given (i)′. To prove (12.9) we use the trivial inequality

$$x^\theta y^{1-\theta}\leqslant x+y, \quad \text{if } 0\leqslant\theta\leqslant 1.$$

Then put $x=A$, $y=A^aB^{1-a}$, and choose θ such that $\theta+a(1-\theta)=a_1$. *q. e. d.*

In defining integral operators on $M=\tilde{M}$, we must first choose a suitable measure. Since the invariant measure on N_F may be given by defining the Riemannian metric which makes $\{Y_{jk}\}_{j,k}$ orthonormal, we impose on $\tilde{M}$ the metric making $\{\tilde{X}_{jk}\}_{j,k}$ orthonormal. For $\tilde{\xi}$ fixed, the differential $d\Theta_{\tilde{\xi}}$ satisfies

$$d\Theta_{\tilde{\xi}}(\tilde{X}_{jk}|_{\tilde{\xi}}) = Y_{jk}|_0.$$

Therefore the volume element $d\tilde{\eta}$ at $\tilde{\xi}$ in $\tilde{M}$ is carried into the volume element du on N_F at 0. Since any smooth function on N_F which vanishes at $u=0$ is $O(|u|^1)$, we have

$$d\tilde{\eta} = (1+O(|u|^1))du. \tag{12.10}$$

Part III. Operators corresponding to free vector fields

We shall begin by giving an outline of the material presented in this and the remaining part of the paper.

We are concerned with the situation that arises when we are given vector fields $X_1, X_2, \ldots, X_n$, on a manifold M, whose commutators (up to length r) span the tangent space at each point. In the previous part we have shown how, in terms of these vector fields, we can embed M locally as a submanifold of a larger $\tilde{M}$ and extend the vector fields to $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_n$ so that these are free (up to step r) and span the tangent space of $\tilde{M}$. Part III will be devoted to the resulting analysis on $\tilde{M}$, and in particular to the study of the integral operators which arise when inverting differential operators of the type

$$\tilde{\mathcal{L}} = \sum_{j=1}^{n} \tilde{X}_j^2 + \tfrac{1}{2}\sum_{j,k} c_{jk}[\tilde{X}_j, \tilde{X}^k].$$

In Part IV we shall apply our results to the original situation involving the vector fields $X_1, X_2, \ldots, X_n$.

To simplify notation we shall in all of Part III omit the $\sim$ and write $X_1, X_2, \ldots, X_n$ for a set of vector fields which are free (up to step r) on M, and whose commutators up to step r span the tangent space at each point. Shrinking M if necessary we obtain from Part II the existence of the important mapping $\Theta: M \times M \to N$, where N is the free nilpotent group on n generators, of step r; (see § 7, Theorem 5).

We shall use Θ to construct our basic integral operators. Our problem then becomes that of proving that these operators satisfy certain commutation relations with differential operators, and that they are bounded on appropriate spaces such as L^2, L^p, L^p_k, S^p_k, and Λ_α. It is here that matters become very similar to the situation already studied in Folland-Stein [8], where the group N was the Heisenberg group. We shall describe the arguments required in full detail only when they are substantially different from those given in that paper. The proofs that can be given merely by paraphrasing similar proofs of [8] will be outlined only. As in § 5 and § 6, we shall assume, however, that our functions may be vector-valued. As before, by definition a vector-valued function will be said to be in a certain class if each of its scalar-valued components is in that class.

§ 13. L^p inequalities for operators of type λ.

In the standard theory of singular integral operators (using the Mihlin-Calderón-Zygmund formalism) the operators are given in the form

$$f \to \int k(x, x-y)\, f(y)\, dy, \tag{13.1}$$

where, for each fixed x, $k(x, z)$ is a homogeneous function in z of suitable degree, and the dependence on x is smooth.

In our situation matters will be similar. The analogue of (13.1) will be

$$f \to \int k_\xi(\Theta(\eta, \xi))\, f(\eta)\, d\eta, \tag{13.2}$$

where each $k_\xi(u)$ will be a homogeneous function of u defined on N, where the homogeity is taken with respect to the standard one appropriate for N. Furthermore, the dependence on ξ will be smooth. Observe that $\Theta(\eta, \xi)$ plays the role of $x-y$.

Before giving formal definitions, we recall some notation concerning groups. $N = N_F$ is defined in § 3. With the standard homogeneity δ_t given on N_F as in § 5, the homogeneous dimension Q is that integer so that $\delta_t(du) = t^Q du$. [1] A function k defined on $N - \{0\}$ which is smooth away from the origin is said to be of *type* λ, $\lambda > 0$, if

$$k(\delta_t(u)) = t^{-Q+\lambda} k(u) \quad \text{all } t > 0.$$

k is said to be of *type* 0 if the above holds with $\lambda = 0$ and the mean value of k vanishes, i.e.

$$\int_{a \leq |u| \leq b} k(u)\, du = 0,$$

where $|\cdot|$ is the norm function. (See (5.2).)

We may assume (replacing M by an open relatively compact subset of M) that the mapping Θ is uniformly smooth on $M \times M$. Write $d\xi$ for the measure given by the Riemannian metric as described in § 12. A function $K(\xi, \eta)$ on $M \times M$ will be said to be a *kernel* of *type* λ, $\lambda \geq 0$ if for any positive integer l we can write

$$K(\xi, \eta) = \sum_{i=1}^{s} a_i(\xi)\, k_\xi^{(i)}(\Theta(\eta, \xi))\, b_i(\eta) + E_l(\xi, \eta). \tag{13.3}$$

---

[1] Footnote: The reader should note that in the case treated in Folland-Stein [8], $Q = 2n + 2$.

(a) $E_l \in C^l(M \times M)$,

(b) $a_i, b_i \in C_0^\infty(M)$, $\quad i = 1, 2, \ldots, s$,

(c) For each i, the function $u \to k_\xi^{(i)}(u)$, defined on N is of type $\geqslant \lambda$ and depends smoothly on ξ.

(Of course the integer s depends on l.)

It is instructive to point out that the definition (13.3) is not as unsymmetric as it might seem. Indeed we claim that if $K(\xi, \eta)$ is a kernel of type λ then so is $K(\eta, \xi)$. To see this, note first that since $\Theta(\xi, \eta) = -\Theta(\eta, \xi)$ we are easily reduced to the statement that if $k_\xi(u)$ is of type λ and smooth in ξ, then $a(\xi) k_\eta(\Theta(\eta, \xi)) b(\eta)$ is a kernel of type λ.

In fact, write $u = (u_{jk}) \in N$, where $u_{jk} = u_{jk}(\eta, \xi)$, and expand the function $\eta \to k_\eta(z)$ in a Taylor expansion about the point ξ in powers of the u_{jk}. Then

$$k_\eta(z) = k_\xi(z) + \sum_{|\alpha|=1}^{t} k_\xi^{(\alpha)}(z) \frac{u^\alpha}{\alpha!} + R_t(z, u),$$

where $k_\xi^{(\alpha)}(z)$ denotes the appropriate partial derivative of $k_\xi(z)$ with respect to ξ, and u^α is the corresponding monomial in the u_{jk}. Observe that if $k_\xi(z)$ is of type λ (as a function of z) then so is $k_\xi^{(\alpha)}(z)$. With $z = (u_{jk}) = \Theta(\eta, \xi)$ the $k_\xi^{(\alpha)}(u) u^\alpha$ are therefore of type $> \lambda$. Finally, if t is sufficiently large, the function $R_t(z, u)$ with $z = u = \Theta(\eta, \xi)$ belongs to $C^l(M \times M)$. This proves the assertion concerning the symmetry of kernels of type λ.

Our final, and basic, definition is now as follows. An operator T is said to be of *type* λ, $\lambda > 0$ if

$$(13.4) \qquad (Tf)(\xi) = \int_M K(\xi, \eta) f(\eta)\, d\eta,$$

where K is a kernel of type λ. We shall show below that this integral converges absolutely. When $\lambda = 0$, we say T is of type λ if

$$(13.5) \qquad Tf(\xi) = \int_M K(\xi, \eta) f(\eta)\, d\eta + a(\xi) f(\xi),$$

where K is of type 0 and $a \in C_0^\infty(M)$. The integral (13.5) will be taken in the principal-value sense.

Several facts will be needed to show that the operators are well-defined. First observe that if K is a kernel of type λ,

$$(13.6) \qquad |K(\xi, \eta)| \leqslant C(\varrho(\xi, \eta))^{-Q+\lambda},$$

where ϱ is the pseudo-metric given by $\varrho(\xi,\eta)=|\Theta(\xi,\eta)|$. Next, by (13.6) and the fact that $d\eta=(1+O(|u|))du$, (see 12.10)

$$\int_{\varrho(\xi,\eta)<C}|K(\xi,\eta)|\,d\eta\leqslant C\int_{|u|\leqslant C}|u|^{-Q+\lambda}du<\infty$$

if $\lambda>0$.(1) Now suppose $\lambda=0$. Using the vanishing mean-value property we have (as on p. 479 of [8])

$$\left|\int_{\varepsilon<\varrho(\xi,\eta)<\delta}K(\xi,\eta)\,d\eta\right|\leqslant C(\delta-\varepsilon) \tag{13.7}$$

from which the existence of the principal value integral defining (13.5) follows easily for smooth f, by using (13.6) (See [8], p. 480).

The first main theorem is as follows. Suppose $\lambda=0$.

For each $\varepsilon>0$ define T_ε by

$$T_\varepsilon(f)(\xi)=\int_{\varrho(\xi,\eta)>\varepsilon}K(\xi,\eta)f(\eta)\,d\eta+a(\xi)f(\xi).$$

THEOREM 6. *Suppose T is an operator of type 0. Then T can be extended to $L^p(M)$, $1<p<\infty$, as follows: For each $f\in L^p$, $\lim_{\varepsilon\to0}T_\varepsilon(f)=T(f)$ exists in the L^p norm, and the mapping $T\to T(f)$ is bounded from L^p to itself.*

Proof. The first step is to prove that T has the required property when $p=2$. For this purpose we may assume $a\equiv0$ in (13.5) and define

$$T_k(f)=\int_{2^{-k}<\varrho(\xi,\eta)\leqslant2^{-k+1}}K(\xi,\eta)f(\eta)\,d\eta,\quad k=0,1,2,\dots.$$

Following the argument of [8], § 15, it suffices to show that

$$\|T_k^*T_j\|\leqslant C2^{-|k-j|/r},$$

and

$$\|T_kT_j^*\|\leqslant C2^{-|k-j|/r}.$$

For simplicity we may assume that

$$K(\xi,\eta)=k_\xi(\Theta(\eta,\xi)),$$

where $k_\xi(u)$ is of type 0.

(1) Footnote: It is useful to recall that $\int_{|u|\leqslant a}|u|^{-Q+\lambda}du=C_\lambda a^\lambda$ if $\lambda>0$ $\int_{|u|\geqslant a}|u|^{-Q+\lambda}du=C_\lambda a^\lambda$ if $\lambda<0$. See Knapp–Stein [12], Folland–Stein [8], Folland [6].

The main ingredients for the proof are:

$$(13.8)\quad \begin{cases} \text{(i)}\ |K(\xi,\eta)| \leqslant A\varrho(\xi,\eta)^{-Q}, \text{ obtained from (13.6) with } \lambda=0; \\ \text{(ii) the mean-value estimate (13.7);} \\ \text{(iii) the estimate} \end{cases}$$

$$|K(\eta,\zeta)-K(\eta,\xi)| \leqslant C_0(\varrho(\xi,\zeta)+\varrho(\xi,\zeta)^{1/r}\varrho(\eta,\zeta)^{1-(1/r)})\varrho(\eta,\zeta)^{-Q-}$$

when $\varrho(\eta,\zeta)>C\varrho(\xi,\zeta)$ for C sufficiently large.

To prove (iii), first write $u=\Theta(\eta,\zeta)$, $v=\Theta(\eta,\xi)$ and note that if $\varrho(\eta,\zeta)>C\varrho(\xi,\zeta)$ then $|u|=\varrho(\eta,\zeta)\geqslant C_1(\varrho(\xi,\zeta)+\varrho(\xi,\zeta)^{1/r}\varrho(\eta,\zeta)^{1-(1/r)}\geqslant C_2|\Theta(\eta,\zeta)-\Theta(\eta,\xi)|$, by (12.3), $=C_2|u-v|$. By choosing C sufficiently large, we may assure that $C_2\geqslant 2$. Therefore, estimating $|K(\eta,\zeta)-K(\eta,\xi)|$, $\varrho(\eta,\zeta)>C\varrho(\xi,\zeta)$ becomes a question of estimating

$$k_\zeta(u)-k_\xi(v) \quad \text{with } |u|\geqslant 2|u-v|. \tag{13.9}$$

Now the difference in (13.9) can be written as

$$\{k_\zeta(u)-k_\xi(u)\}+\{k_\xi(u)-k_\xi(v)\}. \tag{13.10}$$

For the first bracket we have the successive estimates

$$\begin{aligned} |k_\zeta(u)-k_\xi(u)| &\leqslant C\|\zeta-\xi\|\,|k'(u)| \leqslant C\|\zeta-\xi\|\,|u|^{-Q} \\ &\leqslant C|\Theta(\eta,\zeta)-\Theta(\eta,\xi)|\,|u|^{-Q} \leqslant C|\Theta(\eta,\zeta)-\Theta(\eta,\xi)|\,\varrho(\eta,\zeta)^{-Q-1}. \end{aligned}$$

The first estimate above follows from the mean value theorem, the next by the homogeneity of k', the third from the fact that $\zeta\to\Theta(\eta,\zeta)$ is a diffeomorphism, and the last since $|u|=\varrho(\eta,\zeta)\leqslant$ constant. For the second bracket of (13.10), observe that whenever k is a smooth homogeneous function of degree $-Q$, then

$$|k_\xi(u)-k_\xi(v)|\leqslant C\frac{|u-v|}{|u|^{Q+1}} \quad \text{for } |u|\geqslant 2|u-v|,$$

as a simple homogeneity argument shows. (See Lemma 8.10 in [8].) Putting these two together gives

$$|K(\eta,\zeta)-K(\eta,\xi)|\leqslant C|\Theta(\eta,\zeta)-\Theta(\eta,\xi)|\,\varrho(\eta,\zeta)^{-Q-1},$$

from which (13.8 iii) now follows by (12.3).

By using (13.8) the rest of the proof of the L^2 boundedness may be completed as in the case treated in [8]. To prove that T extends to a bounded operator on L^p, $1<p\leqslant 2$, we need to show that for sufficiently large C

$$\int_{\varrho(\xi,\zeta)>C\varrho(\xi,\eta)} |K_\varepsilon(\zeta,\xi)-K_\varepsilon(\zeta,\eta)|\,d\zeta\leqslant C_0,$$

where

$$K_\varepsilon(\zeta, \xi) = \begin{cases} K(\zeta, \xi) & \text{when} \quad \varrho(\zeta, \xi) > \varepsilon \\ 0 & \text{otherwise.} \end{cases}$$

This is proved by the same argument as in [8], upon using (13.8 iii) and the fact that $|K(\xi, \eta)| \leqslant C\varrho(\xi, \eta)^{-Q}$. Finally, the result for $2 \leqslant p < \infty$ then follows from the case $1 < p \leqslant 2$ by duality, in view of the essential symmetry of the kernel $K(\xi, \eta)$ already discussed above.

The L^p theory for operators of type λ, $\lambda > 0$ is much simpler.

THEOREM 7. *Suppose T is of type $\lambda > 0$. Then T has a bounded extension on $L_p(M)$, $1 \leqslant p \leqslant \infty$.* If $0 < \lambda < Q$, *T is bounded from $L^p(M)$ to $L^q(M)$, where $1/q = (1/p) - (\lambda/Q)$, if $1 < p, q < \infty$.*

The proof is not different from the proof of the corresponding theorem in [8]. See in particular Theorem 15.11 and the general lemmas (15.2) and (15.3) used to prove it.

§ 14. Operators of type λ and vector fields

We shall next study the interplay of the basic vector fields $X_1, X_2, \ldots, X_n$ and the operators T of type λ.

Recall (from the definitions of § 7) that we have a doubly indexed family of vector fields $\{X_{jk}\}$, $1 \leqslant j \leqslant r$, where $X_{1k} = X_k$, and more generally the j index indicates that X_{jk} is a commutator of length j of the $X_1, X_2, \ldots, X_n$. The $\{X_{jk}\}$ span the tangent space at each point.

The results which we shall need are as follows.

THEOREM 8. *Suppose T is an operator of type $\lambda \geqslant 1$. Then $X_k T$ and TX_k are operators of type $\lambda - 1$.*[1]

THEOREM 9. *Suppose T is an operator of type $\lambda \geqslant 0$. Given the vector field $X_{j_0 k_0}$, $1 \leqslant j_0 \leqslant r$, there exist operators $\{T_{j,k}\}$ and T_0 so that*

$$X_{j_0, k_0} T = \sum_{j,k} T_{jk} X_{jk} + T_0,$$

where the operator T_{jk} is of type $\lambda + j - j_0$ if $j \geqslant j_0$ or of type λ if $j < j_0$, and T_0 is of type λ.

Before proceding to the proof of Theorems 8 and 9, we review some notions concerning homogeneity with respect to the family of dilations δ_t. Recall that a function f on N

(1) Strictly speaking one should say that there exists an operator T' of type $\lambda - 1$ so that $TXf = T'f$ for smooth functions of compact support. A similar interpretation should be given to all further statements of the same form.

is said to be *homogeneous* of *degree* λ_1 if $\delta_t f(u) = t^{\lambda_1} f(u)$, where $\delta_t f(u) = f(\delta_t u)$. A differential operator D is homogeneous of degree λ_2 if $D(\delta_t h) = t^{\lambda_2} Dh o \delta_t$ for all functions h. Since we are now primarily interested in local questions, we shall say that a function f is of (local) degree $\geqslant \lambda_1$ if its Taylor expansion at 0 is a formal sum of terms each of which is homogeneous of degree $\geqslant \lambda_1$. Similarly, D is of degree $\leqslant \lambda_2$ if its Taylor expansion is a formal sum operators of degrees $\leqslant \lambda_2$. (See § 7.) In the sequel we shall use inplicitly the following simple relations between functions and differential operators:

If f is a function of degree $\geqslant \lambda_1$ and D is a differential operator of degree $\leqslant \lambda_2$, then

(i) fD is a differential operator of degree $\leqslant \lambda_2 - \lambda_1$, and

(ii) Df is a function of degree $\geqslant \lambda_1 - \lambda_2$.

Proof of Theorem 8. We consider first the case when $\lambda > 1$. Using the reasoning in [8], pp. 488–490, it suffices to prove that whenever $K(\xi, \eta)$ is a kernel of type λ, $\lambda > 1$, then $X_k^\xi K(\xi, \eta)$ and $X_k^\eta K(\xi, \eta)$ are kernels of type $\lambda - 1$, where the superscripts indicate the variable of differentiation.

Recall from § 7 that the vector fields X_k can be suitably approximated by the corresponding vector fields Y_k on the free nilpotent Lie group N_F. Now if η is fixed, we express ξ near η in terms of u by $\xi \leftrightarrow u = \Theta(\eta, \xi)$. Then by Theorem 5,

$$X_k^\xi = X_k = Y_k + R_k, \tag{14.3}$$

where R_k is a differential operator of degree $\leqslant 0$.

It will also be useful to express Y_k in canonical coordinates. Since $\delta_t(\exp \tau Y_k) = t \exp \tau Y_k$, Y_k is a homogeneous differential operator of degree 1. Therefore, if we write

$$Y_k = \frac{\partial}{\partial u_{1k}} + \sum_{1 < l \leqslant r} g_{ls}^{(k)}(u) \frac{\partial}{\partial u_{ls}}, \tag{14.4}$$

each $g_{l,s}^{(k)}(u)\partial/\partial u_{l,s}$ is homogeneous of degree 1, and hence $g_{l,s}^{(k)}$ is homogeneous of degree $l-1$.

We are now ready to calculate $X_l^\xi K(\xi, \eta)$. It suffices to restrict attention to a kernel $K(\xi, \eta)$ of the form

$$K(\xi, \eta) = a(\xi) k_\xi(\Theta(\eta, \xi)) b(\eta),$$

where $k_\xi(u)$ is homogeneous of degree $-Q + \lambda$ and depends smoothly on ξ. Then a differentiation on $a(\xi)$ gives a kernel of type λ. Similarly the differentiation with respect to ξ on $k_\xi(u)$ again gives a function of the same type, λ, and this again leads to a kernel of type λ.

Finally we apply X_l^ξ to the ξ variable of $\Theta(\eta, \xi)$ by using (14.3) and (14.4). Observe that since Y_l is a homogeneous differential operator of degree 1, $Y_l(k_\xi(u))$ is homogeneous function $k^1_\xi(u)$ of degree $\lambda-1$, while $R_l(k_\xi(u))$ is of degree $\geqslant\lambda$. If the Taylor expansion of $R_l(k_\xi(u))$ at $u=0$ is taken to enough terms the remainders will contribute a function of u which is of class C^q in u, whenever q is fixed beforehand.

There remarks show that the kernel $X^\xi_i K(\xi,\eta)$ satisfies the expansion property (13.3) for $\lambda-1$ and hence is a kernel of type $\lambda-1$. The proof is the same for $X^\eta_i K(\xi,\eta)$. This concludes the proof of the theorem when $\lambda>1$.

The proof for $\lambda=1$ is similar except for one modification. We need to show that if $k(u)$ is a kernel of type 1 on N_F, i.e. $k(u)$ homogeneous of degree $-Q+1$ and smooth away from the origin, then $Y_{lk}k(u)$ is a homogeneous function of degree $-Q$ with *vanishing mean value*. The latter fact of course follows because $Y_{lk}k$ is a homogeneous distribution of type 0, and all such are, modulo a constant multiple of the dirac delta function, given by a homogeneous function of degree $-Q$ with vanishing mean value. (See e.g. Folland [6], Proposition 1.8.) Now the rest of the proof for the case $\lambda=1$ is completed as in case of $\lambda>1$.

Proof of Theorem 9. We first extend (14.3) and (14.4) to higher commutators. First note that if D_1 and D_2 are differential operators of local degrees $\leqslant\lambda_1$ and $\leqslant\lambda_2$, respectively, then D_1D_2, and therefore $[D_1, D_2]$, are of degree $\leqslant\lambda_1+\lambda_2$. Since each $Y_{j_0k_0}$ is a commutator of length j of the Y_{lk}, and $X_{j_0k_0}$ is the corresponding commutator of the X_{jk}, (14.3) may be generalized to

$$X^\xi_{jk} = Y_{jk}+R_{jk}, \tag{14.3'}$$

where R_{jk} is a differential operator of degree $\leqslant j-1$. (14.4) generalizes to

$$Y_{jk}=\frac{\partial}{\partial u_{jk}}+\sum_{j<l\leqslant r} g^{jk}_{ls}(u)\frac{\partial}{\partial u_{ls}}, \tag{14.4'}$$

where g^{jk}_{ls} is homogeneous of degree $l-j$. This follows from the fact that Y_{jk} is homogeneous of degree j.

We next claim that if R is any vector field of degree $\leqslant j_0$, then

$$R=\sum h_{jk}X^\xi_{jk}, \tag{14.5}$$

where each h_{jk} has a Taylor expansion whose homogeneous terms are functions of degrees $\geqslant j-j_0$. In fact, expressing R first in terms of the $\partial/\partial u_{jk}$, we have

$$R=\sum a_{jk}\frac{\partial}{\partial u_{jk}},$$

where the a_{jk} involve homogeneities of degrees $\geqslant j-j_0$. By inverting the triangular system (14.4), we obtain

$$\frac{\partial}{\partial u_{jk}}= Y_{jk}+\sum_{j<l\leqslant r} f^{(j,k)}_{ls}(u)\, Y_{ls}, \tag{14.6}$$

with each $f_{ls}^{(j,k)}(u)$ homogeneous of degree $l-j$. Substituting this in the expression above for R gives

$$R=\sum b_{jk}Y_{jk},$$

where again the homogeneities of the b_{jk} are of degrees $\geqslant j-j_0$. However since $Y_{jk}=X^\xi_{jk}-R_{jk}$, R_{jk} of degree $\leqslant j-1$ by (14.3′) we may write

$$R=\sum b_{jk}X^\xi_{jk}-\sum b_{jk}R_{jk}.$$

The term $\sum b_{jk}X^\xi_{jk}$ is already in the proper form, while $\sum b_{jk}R_{jk}$ is of degree $\leqslant j_0-1$. Thus we may apply the argument again to $R'=-\sum b_{jk}R_{jk}$. Continuing this way we establish (14.5).

Let us next define the vector fields Y'_{jk} by

$$Y'_{jk}(f)(-u)=-Y_{jk}(f(-u)).$$

Because of (14.6) we have

$$\frac{\partial}{\partial u_{jk}}=Y'_{jk}+\sum_{j<l\leqslant r}f_{l,s}^{(j,k)}(-u)\,Y'_{l,s}$$

and hence by (14.4)

$$Y_{jk}=\sum g_{ls}^{jk}(u)\,Y'_{ls}, \tag{14.7}$$

where g_{ls}^{jk} is homogeneous of degree $l-j_0$. Also, since $\Theta(\xi,\eta)=-\Theta(\eta,\xi)$ it follows from (14.3′) that

$$X^\eta_{jk}=-Y'_{jk}+R'_{jk}, \tag{14.3″}$$

where R'_{jk} is of degree $\leqslant j-1$. Thus if R is any vector field of degree $\leqslant j_0$, then

$$R=\sum h'_{jk}X^\eta_{jk}, \tag{14.5′}$$

where each h'_{jk} has a Taylor expansion of homogeneous functions of degree $\geqslant j-j_0$.

In proving Theorem 9, we need to apply the vector fields $X^\xi_{j_0k_0}$ to terms of the form $k_\xi(\Theta(\eta,\xi))$ and to compare this with the effect of the X^η_{jk} on $k_\xi(\Theta(\eta,\xi))$. When the differentiation is with respect to ξ in $k_\xi(u)$ we get a term of type λ which is incorporated in T_0. Now let X^ξ_{jk} and X^η_{jk} act on the u variable. By (14.3′), (14.3″), (14.5), (14.5′), and (14.7),

$$X^\xi_{j_0k_0}=\sum a_{jk}^{j_0k_0}(u)\,X^\eta_{jk},$$

where each $a_{jk}^{j_0k_0}(u)$ is of local degree $\geqslant \max\{j-j_0,0\}$. Hence,

$$X^\xi_{j_0k_0}(k_\xi(u))=\sum a_{jk}^{j_0k_0}(u)\,X^\eta_{jk}(k_\xi(u))=\sum X^\eta_{jk}(a_{jk}^{j_0k_0}(u)\,k_\xi(u))-\sum X^\eta_{jk}(a_{jk}^{j_0k_0}(u))\,k_\xi(u).$$

$a_{jk}^{j_0k_0}(u)k_\xi(u)$ is a function of type $\geqslant\max\{\lambda,(j-j_0)+\lambda\}$. Thus the operator T_{jk} whose kernel is $K(\xi,\eta)=a(\xi)a_{jk}^{j_0k_0}(u)k_\xi(u)b(\eta)$, with $u=\Theta(\eta,\xi)$, is of type $\max\{\lambda,(j-j_0)+\lambda\}$, and is in the desired form for (14.1). The terms $X^{j}_{j,k}(a^{j_0 k_1}_{j,k}(u))\,k_\xi(u)$ give operators of type λ, which again are incorporated in T_0. This concludes the proof of the theorem.

COROLLARY. *Suppose T is an operator of type λ, $\lambda \geq 0$. Then there exist operators $T_0, T_1, \ldots, T_n$ of type λ so that*

$$X_{k_0} T = \sum_{k=1}^{n} T_k X_k + T_0. \tag{14.8}$$

Proof. Applying Theorem 9, we obtain

$$X_{k_0} T = \sum T_{jk} X_{jk} + T_0,$$

with T_{jk} of type $\lambda + j - 1$. Now consider, for example, the monomial $T_{2k} X_{2k}$, writing $X_{2k} = [X_{1k'}, X_{1k''}]$,

$$T_{2k} X_{2k} = (T_{2k} X_{1k'}) X_{1k''} - (T_{2k} X_{1k''}) X_{1k'}.$$

By Theorem 9, $T_{2k} X_{1k'}$ and $T_{2k} X_{1k''}$ are type λ, since T_{2k} is of type $\lambda + 1$. Hence $T_{2k} X_{2k}$ has the required form for (14.8). A similar argument holds for $T_{jk} X_{jk}$, $j > 2$.

§ 15. Parametrices

We are now ready to achieve one of our main goals. We consider the differential operator

$$\mathcal{L} = \sum_{j=1}^{n} X_j^2 + \tfrac{1}{2} \sum_{j,k} c_{jk} [X_j, X_k], \tag{15.1}$$

where $c = (c_{jk})$ is a skew symmetric matrix of smooth functions on M. For further applications the appropriate conditions to impose on the c_{jk} are of two alternative kinds. Recall that the c_{jk} take values in the space of linear transformations of a finite dimensional vector space W. (See § 5). In the scalar case (i.e. when $\dim W = 1$) we shall require

$$\| \operatorname{Im} c(\xi) \| < 1 \quad \text{for each } \xi \in M, \tag{15.2}$$

where $\| \cdot \|$ denotes the operator norm as in Part I.

When the functions are vector-valued (and c_{jk} is a linear transformation) we shall require the following less explicit condition on the c_{jk}.(1) Suppose $Y_1, Y_2, \ldots, Y_n$ are the generators of the free Lie algebra $\mathfrak{N}_{F,r}$. Then for each $\xi \in M$, we require that

$$\sum \| Y_j f \|^2 \leq C |(\mathcal{L}_\xi f, f)|, \quad \sum \| Y_j f \|^2 \leq C |(\mathcal{L}_\xi^* f, f)|, \quad f \in C_0^\infty, \tag{15.2'}$$

for all $\xi \in M$, where $\mathcal{L}_\xi = \sum Y_j^2 + \frac{1}{2} \sum c_{jk}(\xi) [Y_j, Y_k]$.

(1) Observe by the results of Part I that in the scalar case (15.2) always implies (15.2′), and conversely (15.2′) implies (15.2) when $n \geq 3$ and c is purely imaginary. In the general case it would be interesting to determine more explicit conditions equivalent with (15.2′).

THEOREM 10. *Given* $a \in C_0^\infty(M)$, *there exist operators* P, S, *and* S' *so that*

(i) P *is of type* 2,

(ii) S *and* S' *are of type* 1,

(15.3) (iii) $\mathcal{L}P = aI + S$, $P\mathcal{L} = aI + S'$, *where* I *denotes the identity transformation.*

Proof. We pass to the free nilpotent algebra $\mathfrak{N}$ with generating vector fields $Y_1, Y_2, \ldots, Y_n$ used to approximate the vector fields X_k, by Theorem 5. For each $\xi \in M$ consider the operator

$$\mathcal{L}_\xi = \sum_{j=1}^{n} Y_j^2 + \tfrac{1}{2}\sum c_{jk}(\xi)[Y_j, Y_k].$$

Now if we are in the scalar case and condition (15.2) holds, then so does (15.2′) by (2.11) of Part I. However, if (15.2′) holds in either case, then Theorem 3 of Part I, guarentees that $\mathcal{L}_\xi$ is hypoelliptic and has a unique fundamental solution k_ξ which satisfies the following properties. (See § 5 and § 6).

(1) $u \to k_\xi(u)$ is of type 2 (as a function on N, the group corresponding to $\mathfrak{N}$);

(2) $\mathcal{L}_\xi k_\xi = \delta$, where δ is the Dirac distribution on N;

(3) The function $\xi \to k_\xi(u)$ is smooth in ξ.

We fashion our parametrix out of k_ξ as follows. Consider any C^∞ function b with compact support in M such that $b = 1$ on the support of a. Now let K be the kernel of type 2 given by

$$K(\xi, \eta) = a(\xi)\, k_\xi(\Theta(\eta, \xi))\, b(\eta), \tag{15.4}$$

and P the corresponding type 2 operator

$$f \to \int K(\xi, \eta) f(\eta)\, d\eta.$$

To verify that P satisfies the properties (ii) and (iii) of the Theorem we follow the analogous argument for the proof of Proposition 16.2 of [8], the only difference being that in our case we shall have to carry out an additional differentiation on the ξ variable of $k_\xi(u)$. Thus for fixed η, let us consider the action of differentiation with respect to ξ, given by $\mathcal{L}^\xi(K(\xi, \eta))$. We claim that the result is $a(\xi)\delta_\eta(\xi) + K_1(\xi, \eta)$, where K_1 will be a kernel of type 1 and $\delta_\eta(\xi)$ is the delta function centered at η.

In carrying out the indicated differentiations, it is easy to see that the main contribution arises when all differentiation is on the u variable in $k_\xi(u)$, since the other differentiations will lead to kernels of type 1 or higher. Now we approximate X_k by Y_k with error R_k, R_k of degree $\leqslant 0$. Then

$$\mathcal{L}=\sum_j X_j^2+\tfrac{1}{2}\sum_{jk}c_{jk}[X_j, X_k]=\sum Y_j^2+\tfrac{1}{2}\sum_{j,k}c_{jk}(\xi)[Y_j\, Y_k]+\Delta,$$

where Δ is a differential operator of local degree $\leqslant 1$. The application of $\sum Y_j^2+\frac{1}{2}\sum c_{jk}(\xi)[Y_j, Y_k]$ to $k_\xi(u)$ gives us the delta function at $u=0$ (i.e. at $\xi=\eta$), and thus this term contributes $a(\xi)b(\xi)\delta_\xi(\eta)=a(\xi)\delta_\eta(\xi)$. Since $k_\xi(u)$ is a kernel of type 2, the application of Δ to $k_\xi(u)$ is a kernel of type 1, and so leads to an operator of type 1. Putting the above together shows that $\mathcal{L}P=aI+S$, where S is of type 1. The proof for $P\mathcal{L}$ is similar. q.e.d.

From Theorem 10 we can find a useful expression for f in terms of the $X_k f$.

COROLLARY. *Suppose a is a C^∞ function of compact support. Then there exists operators $T_0, T_1, ..., T_n$, of type 1, such that for any $f\in C_0^\infty(M)$,*

$$af=\sum_{j=1}^{n} T_j X_j f+T_0 f. \tag{15.5}$$

Proof. Apply the theorem to the special case when $\mathcal{L}=\sum_{j=1}^n X_j^2$, and use (15.3). The result is (15.5) with $T_j=PX_j$, $j=1, ..., n$, and $T_0=S'$. Since P is of type 2, $T_j=PX_j$ is of type 1 by Theorem 8 of § 14, and the corollary is proved.

§ 16. The spaces S_k^p, L_α^p, and Λ_α

The basic properties of operators of type λ, and in particular, of the parametrix P and the error terms S and S' of Theorem 10, will be expressed in terms of function spaces which we now study.

For any integer $k\geqslant 0$ and any p, $1<p<\infty$, we define $S_k^p(M)$ to consist of all $f\in L^p(M)$ such that $(X_{i_1}X_{i_2}\, ...\, X_{i_l})f\in L^p(M)$, all $0\leqslant l\leqslant k$. For the norm we take

$$\|f\|_{S_k^p}=\sum_{0\leqslant l\leqslant r}\|X_{i_1}\dots X_{i_l}f\|_{L^p(M)},$$

where the sum is taken over all ordered monomials $X_{i_1}X_{i_2}\, ...\, X_{i_l}$, $0\leqslant l\leqslant r$ of the basic vector fields.

THEOREM 11. *Suppose T is an operator of type λ, where λ is a non-negative integer. Then T is bounded from S_k^p to $S_{k+\lambda}^p$.*

Proof. Write $F = T(f)$. Then successive applications of Theorem 8 and the corollary of Theorem 10 show that

$$X_{i_1} \dots X_{i_{k+\lambda}} F = \sum_{0 \leqslant k' \leqslant k} \sum_{j_1 \dots j_{k'}} T_{j_1 \dots j_{k'}} X_{j_1} \dots X_{j_{k'}} f,$$

where the operators $T_{j_1 \dots k'}$ are each of type 0. By Theorem 6 these operators are bounded on $L^p(M)$ and the assertion of Theorem 11 then follows.

The function spaces $S^p_k(M)$ take into account the special directions given by the basic vector fields $X_1, \dots, X_n$. For some other applications it is useful also to consider the classical Sobolev spaces which don't distinguish these directions. To study these spaces we can embed M (which we may assume has already been shrunk) in $\mathbf{R}^m (m = \dim M)$ by an appropriate coordinate chart. Then any compactly supported function f on M may be extended to all of $\mathbf{R}^m$ by setting it equal to zero outside the coordinate neighborhood corresponding to M. Now let $L^p_\alpha(\mathbf{R}^m)$, $1 < p < \infty$ be the classical Sobolev spaces. (See e.g. Stein [22]). If f is a function on M we shall say $f \in L^p_\alpha(M)$ if $af \in L^p_\alpha(\mathbf{R}^m)$ for every $a \in C^\infty_0(M)$.

Following this convention we shall say that a mapping T taking functions on M to functions on M is bounded from $L^p_{\alpha_1}(M)$ to $L^p_{\alpha_2}(M)$ if for every pair $a, b \in C^\infty_0(M)$, the mapping aTb is bounded from $L^p_{\alpha_1}(\mathbf{R}^m)$ to $L^p_{\alpha_2}(\mathbf{R}^m)$. Of course the bounds of the mappings aTb may depend on the cut-off functions a and b.

THEOREM 12.(1) *Suppose T is an operator of type λ, where λ is a non-negative integer. Then T is a bounded mapping from $L^p_\alpha(M)$ to $L^p_{\alpha+(\lambda/r)}(M)$, for $\alpha \geqslant 0$, $1 < p < \infty$.*

The reader should recall that r is the least integer so that the commutators $[X_{i_1}, [X_{i_2}[\dots X_{i_j}]\dots],]$, $j \leqslant r$, span the tangent space.

The theorem will be based on the following lemma.

(16.1) LEMMA. *Suppose T is an operator of type 1. Then T maps $L^p(M)$ to $L^p_{1/r}(M)$, $1 < p < \infty$.*

The proof of the lemma is merely a reworking of the argument (corresponding to the case $r = 2$) given for Proposition 19.7, pp. 508–514 of Folland-Stein [8], and so we may be brief.

It suffices to consider T with a kernel of the form $a(\xi) k_\xi(\Theta(\eta, \xi)) b(\eta)$, where $k_\xi(u)$ is homogeneous of degree $-Q+1$ and smooth jointly in ξ and u when $u \neq 0$. We replace

(1) An elaboration of the argument shows that the result extends to any non-negative real λ.

$k_\xi(u)$ by the analytic family $zk_\xi(u)|u|^{z-1}$, $0<Re(z)\leqslant r$, and write T_z for the operator with kernel $a(\xi)zk_\xi(\Theta(\eta,\xi))|\Theta(\eta,\xi)|^{z-1}b(\eta)$.

As in [8], and by the use of the reasoning in § 13 and § 14, one can verify the following.

(i) For each $f\in L^p(M)$, $g\in C_0^\infty(M)$, the function $z\to\int T_z(f)g\,d\xi$ is analytic for $0<\mathrm{Re}(z)<r$, continuous in the closure of the strip, and of at most polynomial growth at infinity.

(ii) When $\mathrm{Re}(z)=r$, $T_z(f)\in L_1^p(M)$ and $\|T_zf\|_{L_1^p(\mathbf{R}^m)}\leqslant C(1+|z|)^{r+2}\|f\|_{L^p(M)}$.

(iii) For each z with $\mathrm{Re}(z)=0$, $T_{z'}(f)\to T_z(f)$ in L^p, as $z'\to z$, when $\mathrm{Re}(z')>0$. Moreover,

$$\|T_z(f)\|_{L^p(\mathbf{R}^m)}\leqslant C(1+|z|)^2\|f\|_{L^p(M)}.$$

(In (ii) and (iii) above we have considered the natural extensions of T_z to functions on $\mathbf{R}^m$ as described above.) A combination of (ii) and (iii) via the Calderón interpolation theorem (see [4]) then gives

$$\|T_1(f)\|_{L^p_{1/r}(\mathbf{R}^m)}=\|T(f)\|_{L^p_{1/r}(\mathbf{R}^m)}\leqslant C\|f\|_{L^p(M)}$$

and the lemma is proved.

Still assuming T is of type 1, we shall show that T maps $L_1^p(M)$ to $L^p_{1+(1/r)}(M)$. Consider the vector fields X_{jk}, $1\leqslant j\leqslant r$ spanning the full tangent space. By Theorem 9, if $b\in C_0^\infty(M)$, then

$$X_{j_0k_0}T(bf)=\sum T_{jk}(X_{jk}bf)+T_0bf,$$

where T_{jk} and T_0 are of type 1. If $f\in L_1^p$, each $X_{jk}bf$ is in L^p, and so by the lemma, $X_{j_0k_0}Tbf\in L^p_{1/r}(M)$. Since the X_{jk} span, this implies $aTbf\in L^p_{1+(1/r)}(M)$. By repeating this argument, we can show that T maps $L^p_\alpha(M)$ to $L^p_{\alpha+(1+r)}$ when α is a non-negative integer. The standard interpolation theorem for $L_k^p(\mathbf{R}^m)$ (see Calderón [4]) then shows that the same result holds for any real α, $0<\alpha<\infty$.

Finally, we drop the restriction that the operator T is of type 1. Suppose T is of type $\lambda>1$, and a is any smooth function of compact support. By induction it will suffice to show that

$$aT=\sum T_jT_j', \tag{16.2}$$

where T_j is of type $\lambda-1$ and T_j' is of type 1. To prove (16.2), let P be the parametrix for $\mathcal{L}=\sum X_j^2$. Then

$$aT=TS-\sum(TX_j)(X_jP),$$

which proves (16.2) with $T_0=T$, $T_0'=S$, and $T_j=TX_j$, $T_j'=X_jT$. (The operators are of the required type by Theorem 8.) This concludes the proof of the theorem.

There is the following basic inclusion relating the spaces S_k^p and L_k^p.

THEOREM 13. $S_k^p(M) \subset L_{k/r}^p(M)$.

Proof. Suppose first $f \in S_1^p(M)$. By (15.5)

$$af = \sum_{j=1}^{m} T_j X_j f + T_0 f, \tag{16.3}$$

T_j of type 1. Since $X_j f \in L^p(M)$ by assumption, Theorem 12 shows $af \in L_{1/r}^p(M)$, proving the case $k=1$. For $k=2$ apply (15.5) again to the function af, using another function $a_1 \in C_0^\infty(M)$ with $a_1 = 1$ on the support of a. This gives

$$af = \sum_{j,j'} T_j X_j T'_{j'} X_{j'} f + \sum_j T_j X_j T_0 f.$$

By the corollary to Theorem 9 we can reverse the order of $X_j T'_j$ which gives

$$af = \sum_{j,l,l',j'} T_j T''_l X_{l'} X_{j'} f + \sum_{j,k,l} T_j T'''_k X_l f,$$

where the T_j, T''_l, T'''_k are of type 1. An application of Theorem 12 then shows $f \in L_{2/r}^p(M)$. This argument can be continued indefinitely, giving the theorem.

We now come to the Λ_α spaces. We shall say that f, defined on M, belongs to $\Lambda_\alpha(M)$ if af, extended to $\mathbf{R}^m$, belongs to $\Lambda_\alpha(\mathbf{R}^m)$ for any $a \in C_0^\infty(M)$. For the basic properties of the Λ_α spaces, see e.g. Stein [22]. The main fact we shall prove is the following.

THEOREM 14. *Suppose T is an operator of type λ, $\lambda > 0$. Then*

(a) *T maps $\Lambda_\alpha(M)$ to $\Lambda_{\alpha+(\lambda/r)}(M)$, if $\alpha > 0$,*

(b) *T maps $L^\infty(M)$ to $\Lambda_{\lambda/r}(M)$.*

Remark. For the case $r=2$ see Greiner-Stein [9].

Proof.(1) We consider first the case $0 < \lambda \leq 1$. Everything will be based on the following estimate:

(16.4) LEMMA. *Suppose $K(\xi, \eta)$ is a kernel of type λ, $0 < \lambda \leq 1$. Then*

$$\int_M |K(\xi_1, \eta) - K(\xi_2, \eta)| d\eta \leq C \|\xi_1 - \xi_2\|^{\lambda/r}. \tag{16.5}$$

(The notation $\|\xi_1 - \xi_2\|$ indicates the Euclidean distance between two points $\xi_1, \xi_2 \in M \subset \mathbf{R}^m$.)

(1) We limit ourselves to the case $r \geq 2$. The argument given would have to be modified somewhat to cover the classical case $r=1$ (see also the proof of Lemma 18.3 below).

Proof. We divide the region of integration of the η variable into two subregions:

$$\{\eta: \varrho(\xi_1, \eta) \leqslant d\|\xi_1-\xi_2\|^{1/r}\}$$

and

$$\{\eta: \varrho(\xi_1, \eta) > d\|\xi_1-\xi_2\|^{1/r}\}.$$

The constant d will be chosen in the course of the proof. We can as usual assume that $K(\xi, \eta)=a(\xi)k_\xi(\Theta(\eta, \xi))b(\eta)$, where each $u \to k_\xi(u)$ is a function of type λ on the group and the dependence on ξ is smooth.

Consider first

$$I_1=\int_{\varrho(\xi_1, \eta)\leqslant d\|\xi_1-\xi_2\|^{1/r}} |K(\xi_1, \eta)-K(\xi_2, \eta)|\,d\eta.$$

Observe that $\varrho(\xi_1, \xi_2)=|\Theta(\xi_1, \xi_2)| \leqslant C\|\Theta(\xi_1, \xi_2)\|^{1/r}$ since $\|u\| \leqslant C' \sum_{j=1}^{r} |u_{jk}|^{1/j} \leqslant C\|u\|^{1/r}$. (See (5.2) and 5.4).) However, $\Theta(\xi, \eta)$ vanishes on the diagonal and is jointly smooth. Therefore $\|\Theta(\xi, \eta)\| \leqslant C\|\xi-\eta\|$, which combined with the above gives

$$\varrho(\xi_1, \xi_2) \leqslant C\|\xi_1-\xi_2\|^{1/r}.$$

From this and the "triangle inequality"

$$\varrho(\xi_2, \eta) \leqslant C(\varrho(\xi_1, \eta)+\varrho(\xi_1, \xi_2)).$$

(see (12.3)) we have the containment

$$\{\eta: \varrho(\xi_1, \eta) \leqslant d\|\xi_1-\xi_2\|^{1/r}\} \subset \{\eta: \varrho(\xi_2, \eta) \leqslant C'\|\xi_1-\xi_2\|^{1/r}\}.$$

Hence we get as an estimate for I_1

$$I_1 \leqslant \int_{\varrho(\xi_1, \eta)\leqslant d\|\xi_1-\xi_2\|^{1/r}} |K(\xi_1, \eta)|\,d\eta + \int_{\varrho(\xi_2, \eta)\leqslant C'\|\xi_1-\xi_2\|^{1/r}} |K(\xi_2, \eta)|\,d\eta.$$

In view of (13.6) and the properties of $\Theta(\xi, \eta)$ and the measure $d\eta$ already discussed in § 12, we get

$$I_1 \leqslant C\int_{|u|\leqslant C\|\xi_1-\xi_2\|^{1/r}} |u|^{-Q+\lambda}\,du \leqslant C'\|\xi_1-\xi_2\|^{\lambda/r}$$

when $\lambda>0$. (See Footnote before (13.7).)

Next, consider

$$I_2=\int_{\varrho(\xi_1, \eta)> d\|\xi_1-\xi_2\|^{1/r}} |K(\xi_1, \eta)-K(\xi_2, \eta)|\,d\eta.$$

We shall estimate $|K(\xi_1, \eta)-K(\xi_2, \eta)|$. Except for a trivial term which is dominated by $\|\xi_1-\xi_2\|(\varrho(\xi_1, \eta))^{-Q+\lambda}$, this difference can be estimated by

$$|k_{\xi_1}(u)-k_{\xi_2}(u)| + |k_{\xi_2}(u)-k_{\xi_2}(v)| \tag{16.6}$$

where $u=\Theta(\eta,\xi_1)$ and $v=\Theta(\eta,\xi_2)$. The first difference in (16.6) can be majorized by $C\|\xi_1-\xi_2\|\,|u|^{-Q+\lambda}=C\|\xi_1-\xi_2\|\varrho(\xi_1,\eta)^{-Q+\lambda}$, which again gives a contribution of the size of the trivial term.

We now focus on estimating the main term of (16.6): $|k_{\xi_2}(u)-k_{\xi_2}(v)|$. We write $u\in N$ as $u=(u_{jk})$. Recall that the basic dilations are $(u_{jk})\to(t^j u_{jk})$, $t>0$. We claim that

$$|k_\xi(u)-k_\xi(v)|\leqslant C\sum_{j=1}^{r}\sum_{k}|u_{jk}-v_{jk}|\,|u|^{-Q+\lambda-j}, \tag{16.7}$$

whenever $|u|\geqslant 2|u-v|$. To prove (16.7), we note that since both sides are homogeneous of degree $-Q+\lambda$ under $(u,v)\to(tu,tv)$, it suffices to prove the equality when $|u|=1$ and $|u-v|\leqslant\frac{1}{2}$. In that case (16.7) is an immediate consequence of the smoothness of $k_\xi(u)$ away from $u=0$.

Now $|u-v|=|\Theta(\eta,\xi_1)-\Theta(\eta,\xi_2)|\leqslant C\|\Theta(\eta,\xi_1)-\Theta(\eta,\xi_2)\|^{1/r}\leqslant C_1\|\xi_1-\xi_2\|^{1/r}$

as before. Thus in the region under consideration, $|u|\geqslant d\|\xi_1-\xi_2\|^{1/r}$ implies $|u|\geqslant 2|u-v|$, for $d\geqslant 2C_1$. Now fix d so that $d=2C_1$. Then since $\sum|u_{jk}-v_{jk}|\leqslant C'\|\Theta(\eta,\xi_1)-\Theta(\eta,\xi_2)\|\leqslant C'''\|\xi_1-\xi_2\|$, (16.7) implies

$$|k_\xi(u)-k_\xi(v)|\leqslant C\|\xi_1-\xi_2\|\,|u|^{-Q+\lambda-r}$$

when $|u|\geqslant d\|\xi_1-\xi_2\|^{1/r}$.

Gathering together our estimates gives us

$$|K(\xi_1,\eta)-K(\xi_2,\eta)|\leqslant C\|\xi_1-\xi_2\|(\varrho(\xi_1,\eta))^{-Q+\lambda-r}$$

in the region used to define I_2. Hence we obtain

$$I_2\leqslant C\|\xi_1-\xi_2\|\int_{\varrho(\xi_1,\eta)\geqslant d\|\xi_1-\xi_2\|^{1/r}}\varrho(\xi_1,\eta)^{-Q+\lambda-r}d\eta.$$

However, for any positive s,

$$\int_{\varrho(\xi,\eta)\geqslant s}\varrho(\xi,\eta)^{-Q+\lambda-r}d\eta\leqslant C\int_{|u|\geqslant Cs}|u|^{-Q+\lambda-r}du=C'_{\lambda,r}s^{\lambda-r}$$

provided $\lambda<r$, which holds in our case since $\lambda\leqslant 1$ and $r\geqslant 2$ by assumption. This proves the lemma.

The lemma immediately yields part (b) of the theorem when $0<\lambda\leqslant 1$.

We prove part (a) of the theorem first when $0<\alpha<1$. Suppose $f\in\Lambda_\alpha(\mathbf{R}^m)$. Then there exists a family $f_\varepsilon\in C^\infty(\mathbf{R}^m)$ so that

$$\|(\partial/\partial x_j)f_\varepsilon\|_{L^\infty}\leqslant C\varepsilon^{\alpha-1},\quad j=1,\dots m,\quad\text{and}\quad\|f-f_\varepsilon\|_{L^\infty}\leqslant C\varepsilon^\alpha$$

whenever $0<\varepsilon<\infty$. In fact, we may take $f_\varepsilon(x)=u(x,\varepsilon)$, where u is the Poisson integral of f, as in Chapter V of Stein [22]. The properties above then follow since

$$f-f_\varepsilon=-\int_0^\varepsilon(\partial u/\partial y)\,dy,\qquad \|\partial u/\partial y\|_{L^\infty}\leqslant Cy^{\alpha-1}\quad\text{and}\quad \|\partial u/\partial x_j\|_{L^\infty}\leqslant Cy^{\alpha-1}.$$

Observe $C\leqslant$ constant $\|f\|_{\Lambda_\alpha}$.

Suppose now T is our operator of type $\lambda\leqslant 1$. Write $T'(f)=aT(bf)$ with $a, b\in C_0^\infty(M)$. Let $F=T'(f)$ and $F_\varepsilon=T'(f_\varepsilon)$. Then by the lemma,

$$\|F-F_\varepsilon\|_{\Lambda_{\lambda/r}}\leqslant C\|f-f_\varepsilon\|_{L^\infty}. \tag{16.8}$$

Next note that

$$\|F_\varepsilon\|_{\Lambda_{(\lambda/r)+1}}\leqslant C\left(\|F_\varepsilon\|_{\Lambda_{\lambda/r}}+\sum_{j=1}^m\left\|\frac{\partial}{\partial x_j}F_\varepsilon\right\|_{\Lambda_{\lambda/r}}\right). \tag{16.9}$$

By using Theorem 9, it is an easy matter to express each $(\partial/\partial x_j)\,T'f_\varepsilon$ as

$$\frac{\partial}{\partial x_j}T'(f_\varepsilon)=\sum_{k=1}^m T_{j,k}\left(\frac{\partial}{\partial x_k}f_\varepsilon\right)+T_{j,0}(f_\varepsilon), \tag{16.10}$$

where the $T_{j,k}$ and $T_{j,0}$ are of type λ. Therefore by the lemma and the bound for $\|\partial u/\partial x_j\|_{L^\infty}$, we get $\|(\partial/\partial x_j)F_\varepsilon\|_{\Lambda_{\lambda/r}}\leqslant C\varepsilon^{\alpha-1}$. Combining this with (16.9) gives

$$\|F_\varepsilon\|_{\Lambda_{(\lambda/r)+1}}\leqslant C\varepsilon^{\alpha-1}. \tag{16.11}$$

From (16.8),

$$\|F-F_\varepsilon\|_{\Lambda_{\lambda/r}}\leqslant C\varepsilon^{\alpha}. \tag{16.12}$$

Now we claim that (16.11) and (16.12) show that $F\in\Lambda_{(\lambda/r)+\alpha}$. To see this, we can argue as follows. Since $0<\alpha<1$, $\lambda\leqslant 1$, and $r\geqslant 1$, we have $\lambda/r+\alpha<2$. Therefore, it suffices to show that $|\Delta_h^2F|\leqslant C|h|^{(\lambda/r)+\alpha}$, $|h|\leqslant 1$, where $\Delta_h^2F=F(x+h)+F(x-h)-2F(x)$. Now $\Delta_h^2F=\Delta_h^2(F-F_\varepsilon)+\Delta_h^2(F_\varepsilon)$. However, in general $\|\Delta_h^2G\|_{L^\infty}\leqslant|h|^\beta\|G\|_{\Lambda_\beta}$, if $0<\beta<2$. Therefore, $\|\Delta_h^2F\|_{L^\infty}\leqslant C(|h|^{\lambda/r}\varepsilon^\alpha+|h|^{(\lambda/r)+1}\varepsilon^{\alpha-1}$, by (16.11) and (16.12). We take therefore $\varepsilon=|h|$ and get $\|\Delta_h^2F\|_{L^\infty}\leqslant C|h|^{(\lambda/r)+\alpha}$. This proves that $F\in\Lambda_{(\lambda/r)+\alpha}$ when $0<\alpha<1$. To prove the theorem when $k<\alpha<k+1$, we use the identity (16.10) k times (with f in place of f_ε) and reduce matters to the case $0<\alpha<1$.

Thus the theorem is proved when $0<\lambda\leqslant 1$, and $\alpha>0$ but non-integral. The case when α is an integer then follows by the standard interpolation theorems for the Λ_α spaces. (See O'Neil [18], Taibleson [23], Calderon [4], and Lions-Peetre [17]). Finally, the restriction $0<\lambda\leqslant 1$ is lifted by observing that by (16.2) each operator T of type λ, $1<\lambda$, can be written as a finite sum of operators, where each summand is the product of an operator of type $\lambda-1$ and one of type 1. This completes the proof of Theorem 14.

Part IV. Applications

We present now the main applications of the theory. In § 17 we shall deal with the regularity properties of the operator $\mathcal{L}=\sum_{j=1}^n X_j^2+\frac{1}{2}\sum_{j,k} c_{jk}[X_j, X_k]$, where $X_1 \dots X_n$ are smooth real vector fields on M whose commutators up to step r span the tangent space, and (c_{jk}) is a skew symmetric matrix of smooth functions. As in § 15, we shall allow the c_{jk} to take values in the space of linear transformations of a vector space W. In § 18 we describe how the arguments must be modified to deal with an operator given by $\sum_{j=1}^n X_j^2+X_0$, where now it is assumed that $X_0, \dots, X_n$, and their commutators up to some step span. Finally in § 19 we show how by these methods we can obtain sharp regularity results for the $\Box_b$ Laplacian arising from the $\bar{\partial}_b$ complex on an appropriate C–R manifold. The results for $\bar{\partial}_b$ extend those of Folland-Stein [8] in two ways:

(i) We may allow general metrics,

(ii) The Levi form need not be nondegenerate.

In all our applications we shall lift our initial vector fields $X_1 \dots, X_n$ (and X_0) on M to $\tilde{X}_1, \dots, \tilde{X}_n$ (and $\tilde{X}_0$) on $\tilde{M}$ which are free in the sense described in Part II and in § 18 below. We then use the results of Part III to write down parametrices, estimates, etc. for the corresponding operators on $\tilde{M}$ in place of M. We therefore must begin here by studying the restriction of operators on $\tilde{M}$ to operators on M.

§ 17. Hypoelliptic operators, I: Sum of squares of vector fields

We are given smooth real vector fields on M with the property that commutators of length $\leqslant r$ span the tangent space at every point in some neighborhood of a fixed $\xi_0 \in M$. The construction of Part II allows us to write $\tilde{M}=M\times \mathbf{T}$, (where M has, if necessary, been shrunk to a smaller neighborhood of ξ_0). Recall that in Part III we have written M for $\tilde{M}$. Here $\mathbf{T}$ is the space of additional variables. Shrinking $\tilde{M}$, we will take $\mathbf{T}$ to be an open ball centered around the origin in $\mathbf{R}^{\tilde{m}-m}$, where $\tilde{m}$ is the dimension of the free group $N_{F,r}$ and $m=\dim M$. We may thus assume $\tilde{M}$ is an open subset of $\mathbf{R}^{\tilde{m}}$ with compact closure. The vector fields $X_1, X_2, \dots, X_n$ on M are extended to vector fields $\tilde{X}_1 \dots \tilde{X}_n$ on $\tilde{M}$, which are free up to step r.

We write $\tilde{\xi}=(\xi, t)$ where $\tilde{\xi}\in\tilde{M}$, $\xi\in M$, $t\in\mathbf{T}$. Since M is itself a coordinate neighborhood of some Euclidean space we may take Euclidean measures $d\xi$ and dt on M and $\mathbf{T}$ respectively. As a consequence, the product measure $d\xi dt$ is equivalent with the measure $d\tilde{\xi}$ defined for $\tilde{M}$ in § 12 above, up to multiplication by a factor which is bounded and has bounded inverse. If f is any smooth function on $\tilde{M}$ which is independent of t (i.e. $f(\tilde{\xi})=f(\xi, 0)$) then

$\tilde{X}_i(f) = X_i(f)$, $i = 1, \ldots, n$. In this connection we define the extension operator E which maps functions on M to functions on $\tilde{M}$ in the obvious way:

$$(Ef)(\tilde{\xi}) = f(\xi, t) = f(\xi),$$

where $\tilde{\xi} = (\xi, t) \in \tilde{M}$. We shall also need to define a restriction operator R mapping functions in $\tilde{M}$ to functions on M as follows. Suppose $\varphi \in C_0^\infty(\mathbf{T})$, with $\int_{\mathbf{T}} \varphi(t)\,dt = 1$. Write

$$Rf(\xi) = \int_{\mathbf{T}} f(\xi, t)\,\varphi(t)\,dt. \tag{17.1}$$

The operator R depends on the particular φ we have chosen. We shall keep φ fixed in the rest of the discussion, but it can be observed that our results are largely independent of φ. An obvious relation between E and R is the identity

$$RE = I. \tag{17.2}$$

It is possible to define the function spaces $S_k^p(M)$, $L_\alpha^p(M)$, and $\Lambda_\alpha(M)$ in exactly the same way as the corresponding analogues for $\tilde{M}$ in § 16. (In fact, since in Part III we had used the convention of labeling $\tilde{M}$, $\tilde{X}_i$, etc. by M, X_i, etc., even the notations needed for the present definitions are unchanged.)

The main properties of E and R are contained in the following propositions.

(17.3) PROPOSITION. *E is bounded mapping from $S_k^p(M)$ to $S_k^p(\tilde{M})$; from $L_\alpha^p(M)$ to $L_\alpha^p(\tilde{M})$; and from $\Lambda_\alpha(M)$ to $\Lambda_\alpha(\tilde{M})$.*

(17.4) PROPOSITION. *R is a bounded mapping from $S_k^p(\tilde{M})$ to $S_k^p(M)$; from $L_\alpha^p(M)$ to $L_\alpha^p(M)$; and from $\Lambda_\alpha(\tilde{M})$ to $\Lambda_\alpha(M)$.*

Proof. Suppose $f \in L^p(M)$. To see that $Ef \in L^p(\tilde{M})$ we check that $\int_{\tilde{M}} |f(\xi, t)|^p d\tilde{\xi} < \infty$. However, as we have observed, $d\tilde{\xi} \leq c\,d\xi\,dt$. Thus $\int_{\tilde{M}} |f(\xi, t)|^p d\xi\,dt \leq C \int_{\mathbf{T}} \{\int_M |f(\xi)|^p d\xi\}\,dt$. Since $\mathbf{T}$ is an open ball, it has finite Euclidean measure. Thus the integral on the right above is finite, which shows that $Ef \in L^p(\tilde{M})$.

Next, it is obvious from the definition of E and from the properties of the $\tilde{X}_j$ as extensions of X_j that

$$\tilde{X}_{j_1} \tilde{X}_{j_2} \ldots \tilde{X}_{j_l} E(f) = E X_{j_1} \ldots X_{j_l}(f). \tag{17.5}$$

From this and the previous remarks it is clear that E maps $S_k^p(M)$ boundedly to $S_k^p(\tilde{M})$. To study the mapping E on the spaces L_α^p, recall the definition of L_α^p given in § 16. A function f defined in M is in, say, $L_\alpha^p(M)$ if for every $a \in C_0^\infty(M)$ the function af (regarded as a function on $\mathbf{R}^m$ when extended by zero outside M) belongs to the space $L_\alpha^p(\mathbf{R}^m)$. To prove that E

is bounded, we must show that for every $a \in C^\infty(M)$ and $b \in C_0^\infty(\tilde{M})$, the mapping $f \to bE(af)$ induces a bounded mapping of $L^p_\alpha(\mathbf{R}^m)$ to $L^p_\alpha(\mathbf{R}^{\tilde{m}})$, where $\tilde{X} = \dim \tilde{M}$. Now observe that

$$\left(\frac{\partial}{\partial \xi}\right)^{\gamma_1}\left(\frac{\partial}{\partial t}\right)^{\gamma_2} E(f) = \begin{cases} E\left(\left(\dfrac{\partial}{\partial \xi}\right)^{\gamma_1} f\right) & \text{if } \gamma_2 = 0 \\ 0 & \text{if } \gamma_2 \neq 0. \end{cases} \tag{17.6}$$

From this it follows that E is bounded from L^p_k to $L^p_k(\tilde{M})$ when k is an integer, and the general case for all α positive is then proved by the standard interpolation theorem for L^p_α spaces. (See Calderón [4].)

Finally, it is easy to see from the definitions that E is bounded from $\Lambda_\alpha(M)$ to $\Lambda_\alpha(\tilde{M})$ when $0 < \alpha < 1$. Using (17.6) leads immediately to an extension of this to all positive non-integral values of α. The integral values of α are then obtained by the interpolation theorem for Λ_α spaces. (See the references quoted at the end of § 16.) Proposition 17.3 is then proved.

To prove Proposition 17.4, notice first that $|(Rf)(\xi)| \leq \int_{\mathbf{T}} |f(\xi, t)|\,|\varphi(t)|\,dt$, and so by Hölder's inequality, $|(Rf)(\xi)|^p \leq C \int_{\mathbf{T}} |f(\xi, t)|^p dt$. An integration in ξ gives $\int_M |(Rf)(\xi)|^p d\xi \leq C \int_M \int_{\mathbf{T}} |f(\xi, t)|^p d\xi dt \leq C' \int_{\tilde{M}} |f(\tilde{\xi})|^p d\tilde{\xi}$, which shows that R is a bounded mapping from $L^p(\tilde{M})$ to $L^p(M)$. We may generalize this by replacing $\varphi(t)$ by a function $\lambda(\xi, t) \in C^\infty(\tilde{M})$, which together with its derivatives is bounded on $\tilde{M}$. If R' is defined by

$$(R'f)(\xi) = \int_{\mathbf{T}} \lambda(\xi, t)\, f(\xi, t)\, dt, \tag{17.7}$$

the above argument then shows

$$\|R'f\|_{L^p(M)} \leq C\|f\|_{L^p(\tilde{M})}. \tag{17.8}$$

Next we can see that

$$X_k R(f) = R\tilde{X}_k f + R'(f), \tag{17.9}$$

where R' is of the form (17.7). In fact, by (7.2),

$$X_k R - R\tilde{X}_k = -\sum_l \int_{\mathbf{T}} \lambda_{kl}(\xi, t)\, \varphi(t) \frac{\partial}{\partial t_l} f(\xi, t)\, dt.$$

An integration by parts then gives the desired formula with $\lambda(\xi, t) = \sum_l \partial/\partial t_l (\lambda_{kl}(\xi, t)\varphi(t))$. Repeated application of formulae of the type (17.9) yields

$$X_{k_1} \ldots X_{k_j} R = R\tilde{X}_{k_1} \ldots \tilde{X}_{k_j} + \sum_{\substack{\alpha = (\alpha_1, \ldots, \alpha_l) \\ 0 \leq l < j}} R'_\alpha \tilde{X}_{\alpha_1} \ldots \tilde{X}_{\alpha_l} + R'_0, \tag{17.10}$$

where each R'_α and R'_0 is of the form (17.7).

From (17.10) it is then obvious that R takes $S_k^p(\tilde{M})$ to $S_k^p(M)$. The proof that R takes $L_\alpha^p(\tilde{M})$ to $L_\alpha^p(M)$, and $\Lambda_\alpha(\tilde{M})$ to $\Lambda_\alpha(M)$ is similar to that for E once we observe that

$$\left(\frac{\partial}{\partial\xi}\right)^\gamma R(f)=R\left(\left(\frac{\partial}{\partial\xi}\right)^\gamma f\right).$$

Now Proposition 17.4 is also proved.

For later purposes we observe that in the course of the above proof we have also in effect shown the following.

(17.11) COROLLARY. *If R' is an operator of the form* (17.7), *then R' takes $S_k^p(\tilde{M})$ to $S_k^p(M)$, $L_\alpha^p(\tilde{M})$ to $L_\alpha^p(M)$, $\Lambda_\alpha(\tilde{M})$ to $\Lambda_\alpha(M)$, and $L^\infty(\tilde{M})$ to $L^\infty(M)$.*

Given any operator $\tilde{T}$ mapping functions on $\tilde{M}$ to functions on M we may define an operator T, mapping functions on M to functions on M by

$$T=R\tilde{T}E.$$

An operator T on M will be said to be *smoothing of order* λ (λ is a non-negative integer) if it maps $S_k^p(M)$ to $S_{k+\lambda}^p(M)$, $L_\alpha^p(M)$ to $L_{\alpha+\lambda/r}^p(M)$, $\alpha\geqslant 0$, $\Lambda_\alpha(M)$ to $\Lambda_{\alpha+\lambda/r}(M)$, $\alpha>0$, and $L^\infty(M)$ to $\Lambda_{\lambda/r}(M)$.

(17.12) PROPOSITION. *If $\tilde{T}$ is of type λ, then T is smoothing of order λ. If $\lambda\geqslant 1$, then X_kT and TX_k are smoothing of order $\lambda-1$, $1\leqslant k\leqslant n$.*

The first statement follows from the definitions, Propositions 17.3 and 17.4, and the results in § 16. For the second, use (17.9) and Theorem 11 of § 16.

We are now in a position to obtain our basic results for the parametrix of the operator $\mathcal{L}=\sum_{j=1}^n X_j^2+\frac{1}{2}\sum_{j,k}c_{jk}[X_j,X_k]$ on M. We shall assume that condition (15.2) holds for $\mathcal{L}$ in the scalar case, or that (15.2′) holds if $\mathcal{L}$ is an operator on vector-valued functions.

THEOREM 15. *Given $a\in C_0^\infty(M)$, there exist operators P, S, and S' so that*

$$(17.13)\quad\begin{cases}\text{(a) } P \text{ is smoothing of order } 2,\\ \text{(b) } S \text{ and } S' \text{ are smoothing of order } 1,\\ \mathcal{L}P=aI+S,\ P\mathcal{L}=aI+S'.\end{cases}$$

Proof. Let $\psi\in C_0^\infty(\mathrm{T})$ with the property that $\psi(t)=1$ on the support of φ. (Recall that $(Rf)(\xi)=\int_{\mathrm{T}}f(\xi,t)\varphi(t)dt$.) Write $\tilde{a}(\xi,t)=a(\xi)\psi(t)$. Then according to Theorem 10 in § 15 there exist operators $\tilde{P}$, $\tilde{S}$, and $\tilde{S}_1$ respectively of types 2, 1, and 1, so that

$$\tilde{\mathcal{L}}\tilde{P}=\tilde{a}I+\tilde{S}\quad\text{and}\quad\tilde{P}\tilde{\mathcal{L}}=\tilde{a}I+\tilde{S}_1.$$

Now multiply the second equation by the operator R on the left and by the operator E on the right. It is clear that $R\tilde{a}IE = aI$. Define $S_1 = R\tilde{S}_1E$ and $P = R\tilde{P}E$. By (17.5) we also have $\tilde{\mathcal{L}}E = E\mathcal{L}$. Putting all these together gives $P\mathcal{L} = aI + S_1$, and half of our claim is proved if we invoke proposition 17.12.

Next note that $\mathcal{L}P = \mathcal{L}R\tilde{P}E = R\tilde{\mathcal{L}}\tilde{P}E + [\mathcal{L}R - R\tilde{\mathcal{L}}]\,\tilde{P}E$. However, by (17.10) $\mathcal{L}R - R\tilde{\mathcal{L}} = \sum_{k=1}^{n} R_k'\tilde{X}_k + R_0'$ and so

$$\mathcal{L}P = aI + R\tilde{S}E + \sum_k R_k'\,\tilde{X}_k\tilde{P}E + R_0'\tilde{P}E.$$

Thus we may take $S = R\tilde{S}E + \sum R_k'\tilde{X}_k\,\tilde{P}E + R_0'\,\tilde{P}E$, and the required property for S follows from Corollary 17.11, Proposition 17.3, and Theorem 8 in § 14. The theorem is therefore proved.

By an iteration argument it is possible to refine the parametrix so that the error terms in (17.13) are smoothing of any preassigned order. We shall exhibit a result of this kind in the form most convenient for applications.

(17.14) COROLLARY. *Suppose a, b, c are in $C_0^\infty(M)$ with $b=1$ on the support of a, and $c=1$ on the support of b. Then for each integer l there exist operators P_l and S_l so that*

$$P_l b\mathcal{L} = aI + S_l\cdot c, \tag{17.15}$$

where P_l is smoothing of order 2, and S_l is smoothing of order l.

Proof. We shall use the partial ordering $a \prec b$ to mean that $b=1$ on the support of a. Suppose we are given any three functions α, β, γ in $C_0^\infty(M)$ with the property $\alpha\prec\beta\prec\gamma$. We can always find another function β_0 so that $\alpha\prec\beta_0\prec\beta$. Now apply Theorem 15 with a replaced by α and f replaced by $\beta_0 f$. The result is $P(\mathcal{L}\beta_0) = \alpha I + S'\beta_0$. However, $\mathcal{L}(\beta_0 f) = \beta_0\mathcal{L}(f) + 2\sum_j X_j(\beta_0)X_j(f) + f\mathcal{L}(\beta_0)$. Now put $P_1 = P\beta_0$ and

$$S_1 f = S'(\beta_0 f) - 2\sum_j P(X_j(\beta_0)X_j(f)) - P((\mathcal{L}(\beta_0))f).$$

Since $\beta_0\beta = \beta_0$, it follows that $P_1\beta = P_1$. Also $X_j(\beta_0)\gamma = X_j(\beta_0)$, $(\mathcal{L}\beta_0)\gamma = \mathcal{L}\beta_0$, and $\beta_0\gamma = \beta_0$ imply $S_1\gamma f = S_1 f$. Thus we obtain

$$P_1(\beta\mathcal{L}(f)) = \alpha f + S_1(\gamma f), \tag{17.16}$$

which proves the lemma for $l=1$.

To prove the general case, we proceed by induction. Suppose we wish to prove the result for l, given the triple of functions a, b, c with $a\prec b\prec c$. Choose a', b', c', and α, β, γ with

$$a = a' \prec b' \prec c' \prec \alpha \prec \beta \prec b \prec c \prec \gamma,$$

such that α, β, γ satisfy (17.16). By induction there exist operators P_{l-1} and S_{l-1} so that

$$P_{l-1}(b'\mathcal{L}) = a' + S_{l-1}\cdot c' \tag{17.17}$$

Since $c' \prec \alpha$, then $c'\alpha = c'$. Multiplying (17.16) by c' and substituting for $c'f$ in (17.17), we obtain

$$P_{l-1}(b'\mathcal{L}(f)) = a'f + S_{l-1}(c'(P_1(\beta\mathcal{L}(f)) - c'S_1(\gamma f)) \tag{17.18}$$

Now take $P_l = P_{l-1}(b') - S_{l-1}c'P_1\beta$ and $S_l = S_{l-1}cS_1\gamma$. Since $b'b = b'$, $\beta b = \beta$, $a = a'$, and $\gamma c = \gamma$, we obtain (17.15) from (17.18). The corollary is therefore proved.

THEOREM 16. *Suppose f is locally in $L^p(M)$, and $\mathcal{L}(f) = g$.*

(a) *If $g \in L^p_\alpha(M)$, then $f \in L^p_{\alpha+(2/r)}(M)$, $\alpha \geq 0$.*

(b) *If $g \in \Lambda_\alpha(M)$, then $f \in \Lambda_{\alpha+(2/r)}(M)$, $\alpha > 0$.*

(c) *If $g \in L^\infty(M)$, then $f \in \Lambda_{2/r}(M)$.*

(d) *If $g \in S^p_k(M)$, then $af \in S^p_{k+2}(M)$, for each $a \in C_0^\infty(M)$, $1 < p < \infty$, $k = 0, 1, \ldots$.*

Proof. We are assuming that $\mathcal{L}(f) = g$ holds in the weak sense, i.e. $\int_M f\overline{\mathcal{L}^*(\varphi)}\,d\xi = \int_M g\bar{\varphi}\,d\xi$, for all $\varphi \in C_0^\infty(M)$. Now apply the adjoint of the identity (17.15) and insert in the above. The result is

$$\int_M P_l(bg)\,\bar{\varphi}\,d\xi = \int_M (af + S_l(f))\,\bar{\varphi}\,d\xi,$$

and therefore

$$af = P_l(bg) - S_l(cf) \tag{17.19}$$

This identity holds for all triples a, b, $c \in C_0^\infty(M)$ with $a \prec b \prec c$, and all l. P_l is smoothing of order 2, and S_l is smoothing of order l. Choose $l \geq \alpha + (2/r)$. Then $S_l(cf) \in L^p_{\alpha+(2/r)}(M)$, and $P_l(bg) \in L^p_{\alpha+(2/r)}(M)$. Since a is an arbitrary function in $C_0^\infty(M)$, it follows from (17.19) and the above that $f \in L^p_{\alpha+(2/r)}(M)$. Parts (b), (c), and (d) of Theorem 16 have parallel proofs.

The theorem may be strengthened by assuming merely that f is a distribution on M with $\mathcal{L}(f) = g$, where g is in one of the spaces $L^\infty(M)$, $L^p_\alpha(M)$, $\Lambda_\alpha(M)$, or $S^p_k(M)$. The proof would require an elaboration of the argument, which we shall refrain from giving.

Our last result here is the sharp form of estimates studied in the L^2 context by Hörmander, Kohn, and Radkevitch.

THEOREM 17. *Suppose $X_1, X_2, \ldots, X_n$ are real smooth vector fields whose commutators of length $\leq r$ span the tangent space at each point. If f and $X_i f \in L^p_\alpha(M)$ (resp. $\Lambda_\alpha(M)$) then $f \in L^p_{\alpha+(1/r)}(M)$ (resp. $\Lambda_{\alpha+(1/r)}(M)$) when $1 < p < \infty$ and $\alpha \geq 0$ (resp. $\alpha > 0$).*

Proof. Let $f \in L^p_\alpha(M)$ such that each $X_i f \in L^p_\alpha(M)$ also, and put $\tilde{f} = E(f)$, the extension of f to $\tilde{M}$. By the Corollary of Theorem 10 in § 15 (see (15.5)), there exist operators $\tilde{T}_0, \tilde{T}_1, \ldots, \tilde{T}_n$ of type 1 such that $\tilde{f} = \sum_{i=1}^n \tilde{T}_i \tilde{X}_i(\tilde{f}) + \tilde{T}_0(\tilde{f})$. Now we have already observed that $\tilde{X}_i(Ef) = E(X_i f)$. Hence $\tilde{f} \in L^p_\alpha(\tilde{M})$ and $\tilde{X}_i \tilde{F} \in L^p_\alpha(\tilde{M})$ by Proposition 17.3. Thus by Theorem 12 in § 16, $f \in L^p_{\alpha+(1/r)}(\tilde{M})$, and finally by Proposition 17.4, $f \in L^p_{\alpha+(1/r)}(M)$. The proof for Λ_α is similar.

It is easy to see that the argument also gives the following inequality. Whenever a and b are in $C_0^\infty(M)$ with $b = 1$ on the support of a, then

$$\|af\|_{L^p_{\alpha+(1/r)}(M)} \leqslant C \sum_{j=1}^n (\|bX_j f\|_{L^p_\alpha(M)} + \|bf\|_{L^p_\alpha(M)}). \tag{17.20}$$

Finally, it may be worthwhile to exhibit a parametrix more explicitly. We shall consider the case where $\mathcal{L} = \sum_{j=1}^n X_j^2$. The operator P has a kernel which, apart from some cut-off functions, is given by

$$\int_{\mathbf{T}} \int_{\mathbf{T}} \varphi(t)\, k(\Theta((\xi, t), (\eta, s)))\, dt\, ds, \tag{17.21}$$

$(\xi, \eta) \in M \times M$. Here k is the fundamental solution (kernel) for the operator $\sum_{j=1}^n Y_j^2$ on the free group $N_{F,r}$, as in § 5 of Part I; (ξ, t) and (η, t) are points on the extended manifold $\tilde{M} = M \times \mathbf{T}$; Θ is the basic mapping of $\tilde{M} \times \tilde{M}$ to $N_{F,r}$ described in § 7, and $\varphi \in C_0^\infty(\mathbf{T})$ with $\int_{\mathbf{T}} \varphi(t)\, dt = 1$.

§ 18. Hypoelliptic Operators, II: Operators of Hörmander type

We consider the differential operator $\mathcal{L}$ given by

$$\mathcal{L} = \sum_{j=1}^n X_j^2 + X_0,$$

where $X_0, X_1, \ldots, X_n$ are real smooth vector fields on M; we assume that these vector fields, together with their commutators of some finite order, span the tangent space at any point. We shall show here how the theory which has been worked out in particular for operators of the form $\sum_{j=1}^n X_j^2$ must be modified to take care of this more general case.

First we define the versions of the spaces $S^p_k(M)$ in the present context. Let $X_{j_1} \ldots X_{j_l}$ be a monomial with $0 \leqslant j_s \leqslant n$, $s = 1, \ldots, l$. We shall say that this monomial has *weight* r if $r = r_1 + 2r_2$, where r_1 is the number of X_j's that enter with j between 1 and n, and r_2 is the number of X_0's. So in computing the total weight, we count each $X_1, X_2, \ldots, X_n$ to have weight 1 and X_0 to have weight 2. Similarly, the weight of a commutator $[X_{j_1},$

$[X_{j_1}, [\ldots, X_{j_l}]]\ldots]$ is defined to be equal to the weight of the corresponding monomial $X_{j_1} X_{j_2} \ldots X_{j_l}$.[1]

Now when k is an integer, $S_k^p(M)$ is the collection of all $f \in L^p(M)$ such that $X_{j_1} \ldots X_{j_l} f \in L^p(M)$ for all monomials of weight $\leqslant k$. For the norm we take

$$\|f\|_{S_k^p} = \sum \| X_{j_1} \ldots X_{j_l} f \|_{L^p(M)},$$

where the sum is taken over all ordered monomials of weight $\leqslant k$. Observe that in a sense this definition is not entirely optimal when k is an odd integer. In fact, when $k=1$ the direction X_0 is not explicitly involved. We shall return to this point later.

The spaces $L_\alpha^p(M)$, $\alpha \geqslant 0$, and $\Lambda_\alpha(M)$ are of course defined as before. With these definitions the main regularity result for solution of $\mathcal{L}(f)=g$ is then as follows.

THEOREM 18. *Suppose $\mathcal{L} = \sum_{j=1}^n X_j^2 + X_0$, where all commutators of weight $\leqslant r$ span the tangent space at each point, and $\mathcal{L}(f)=g$, $f \in L^p(M)$, $1<p<\infty$. Then the regularity results for f, in terms of g, given in Theorem* 16 (§ 17) *hold for $\mathcal{L}$ as well.*

In complete analogy with what has been done through § 17, we can proceed to prove Theorem 18 by the following steps.

Step 1. Let $\mathfrak{G}_{F_0}$ be the free Lie algebra on $n+1$ generators $Y_0, Y_1, \ldots, Y_n$ and let $\mathfrak{J}$ be the ideal spanned by all commutators of weight $\geqslant r+1$. (Recall that Y_0 has weight 2, while $Y_1, Y_2, \ldots, Y_n$ each have weight 1.) Then $\mathfrak{N} = \mathfrak{N}_{F,r}^{II} = \mathfrak{G}_{F_0}/\mathfrak{J}$ will be called the *free nilpotent Lie algebra* of type II, on $n+1$ generators and of weight r. We can identify $Y_0, \ldots, Y_n$ with their images in $\mathfrak{N}_{F,r}^{II}$. (We may assume $r \geqslant 2$.) Now the mappings:

$$\delta_t \begin{cases} Y_0 \to t^2 Y_0 \\ Y_j \to t_j Y_j, \quad j=1, \ldots, n \end{cases}$$

$0<t<\infty$, are easily seen to extend to automorphisms of $\mathfrak{N}$. These dilations give a grading of $\mathfrak{N} = \sum_{j=1}^r \oplus V_j$. V_j is the subspace of $\mathfrak{N}$ on which the action of δ_t is given by t^j. Alternatively, V_j is spanned by the commutators of weight j. Thus V_1 is spanned by $Y_1, \ldots, Y_n$, and V_2 is spanned by Y_0 and the $[Y_j, Y_k]$, $1 \leqslant j < k \leqslant n$. $\mathfrak{N}$ is then stratified of type II, according to the definition in § 3.

Now the left-invariant differential operator $\mathcal{L}_0 = \sum_{j+1}^n Y_j^2 + Y_0$ is homogeneous of degree 2 with respect to the above dilations. Furthermore, $\mathcal{L}_0$ and $\mathcal{L}_0^*$ are both hypoelliptic, according to Hörmander's theorem in [10]. Thus by Proposition A in § 5, there exists a unique distribution k of type 2 so that $\mathcal{L}_0(k) = \delta$. This is the fundamental solution of $\mathcal{L}_0$.

[1] Obviously the weight of a commutator is not smaller than its length.

Step 2. Returning to our original vector fields $X_0, X_1, \ldots, X_n$ on M, we shall say that these are free up to weight s, $s=1, 2, \ldots$ at a given point $\xi_0 \in M$ if all commutators of weight $\leqslant s$, restricted to ξ_0, span a subspace of maximal dimension, i.e. of dimension equal to $\dim(\mathfrak{N}^{II}_{F,s})$.

Now if we know that the commutators of weight $\leqslant r$ span the tangent space at a point ξ_0, then in terms of additional variables, we can lift the vector fields $X_0, X_1, \ldots, X_n$ to vector fields $\tilde{X}_0, \tilde{X}_1, \ldots, \tilde{X}_n$ on $\tilde{M}$, so that these vector fields are free up to weight r; moreover, the commutators of weight $\leqslant r$ will span the tangent space at the point $(\xi_0, 0)$; here $\tilde{M}=M\times \mathbf{R}^q$. Secondly, if we define the mapping Θ as in § 7, then at each point $\tilde{\xi}\in\tilde{M}$ we can introduce the canonical coordinate system centered at $\tilde{\xi}$. In terms of local coordinates around any point $\tilde{\eta}$ near ξ we have

$$\tilde{X}_j = Y_j + R_j,$$

where R_j has local degree $\leqslant 0$ if $j=1, \ldots, n$, and R_0 has local degree $\leqslant 1$. (Recall that the variable u_0 dual to Y_0 has degree 2, while the variables u_j dual to Y_j, $j=1, \ldots, n$ have degree 1.)

The proofs of these facts are entirely similar to the situation detailed in Part II.

Step 3. We can now write down a parametrix for $\tilde{\mathcal{L}} = \sum_{j=1}^n \tilde{X}_j^2 + \tilde{X}_0$. It is given by the kernel $K(\xi, \eta) = a(\xi)\,k(\Theta(\eta, \xi))\,b(\eta)$, where k is the fundamental solution for $\sum_{j=1}^n Y^2 + Y_0$ discussed above, and a and $b\in C_0^\infty(\tilde{M})$ with $b=1$ on the support of a.

If we define operators of type λ in the same way as in § 14, then the result of Theorem 10 holds for $\tilde{\mathcal{L}}$ as well; that is, the operator P whose kernel is K (the parametrix) is of type 2, and $\mathcal{L}P - aI$ and $P\mathcal{L} - aI$ are both of type 1.

Step 4. The results of § 16 can be extended, in large part, to include the current case. Thus we can show that if T is an operator of type λ, then if λ is a non-negative integer, T maps S_k^p to $S_{k+\lambda}^p$, and T maps L_α^p to $L^p_{\alpha+(\lambda/r)}$ if $1<p<\infty$ and $0\leqslant\alpha<\infty$. T maps Λ_α to $\Lambda^p_{\alpha+(\lambda/r)}$ and L^∞ to $\Lambda_{\lambda/r}$ if $\alpha>0$ and $\lambda>0$. However in general it is *not* true that $S_k^p\subset L^p_{k/r}$, although inclusion does hold if k is even or a multiple of r. (We have previously alluded to the reason the inclusion may fail in e.g. the case when $k=1$.) The difference from the previous case arises essentially from the fact that the representation formula (15.5) for f in terms of $\tilde{X}_j f$ is no longer available. Instead we must use the formula

$$aI = P\sum_{j=1}^n \tilde{X}_j^2 + P\tilde{X}_0 - S', \tag{18.1}$$

where P is the parametrix for $\sum_{j=1}^n X_j^2 + X_0$, and hence of type 2, and S' is of type 1.

We outline next a proof of the assertions made above about operators of type λ.

(18.2) LEMMA. *If T is of type λ, $\lambda=1$ or 2, then T maps $L^p(\tilde{M})$ to $L^p_{\lambda/r}(\tilde{M})$, $1<p<\infty$.*

The proof is a straightforward modification of the proof of the corresponding lemma, (16.1) in § 16. (The inclusion $S^p_r \subset L^p_1$, needed for the proof of the lemma still holds in this case.)

One also needs the following analogue of Lemma 16.4.

(18.3) LEMMA. *Suppose $K(\xi, \eta,)$ is a kernel of type λ, $0<\lambda\leqslant 2$. Then if $Tf=\int K(\xi,\eta)f(\eta)d\eta$, T maps $L^\infty(\tilde{M})$ to $\Lambda_{\lambda/r}(\tilde{M})$.*

Proof. If we were to exclude the case $r=1$ (which corresponds to the classical case) and also $r=2$, then we would have $\lambda/r<1$, and the argument given for Lemma 16.4 actually gives the desired result without modification. But when $r=2$ or $r=1$, then λ/r may be $\geqslant 1$ and the condition that $Tf\in\Lambda_{\lambda/r}$ is more complicated to verify and thus needs an additional argument. One can proceed as follows. We assume $r=2$ and $\lambda=2$, since the cases $r>2$ or $r=2$, $0<\lambda<2$ have already been taken care of and the case $r=1$ can be treated similarly.

We can restrict consideration to T with kernel of the form $a(\xi)k(\Theta(\eta,\xi))b(\eta)$, where $k(u)$ is homogeneous of degree $-Q+2$. (Q is the homogeneous dimension of the free nilpotent algebra of type II and weight r.) Let $|\cdot|$ denote the norm function, and for complex z in the strip $|\mathrm{Re}\, z|\leqslant 1$, we write T_z for the operator whose kerenel is $a(\xi)k(\Theta(\eta,\xi))\cdot|\Theta(\eta,\xi)|^z b(\eta)$. Using the arguments of the proof of Lemma 16.1, one shows that whenever $f\in L^\infty(M)$ and $g\in L^1(M)$ the following hold.

(i) The function $z\to\int_{\tilde{M}} T_z(f)g\,d\xi$ is analytic and bounded in the strip $-1\leqslant \mathrm{Re}\,(z)\leqslant 1$.

(ii) When $\mathrm{Re}\,(z)=-1$, $T_z(f)\in\Lambda_{1/2}(\tilde{M})$, and $\|T_z f\|_{\Lambda_{1/2}(\tilde{M})}\leqslant A(1+|z|)\|f\|_\infty$.

(iii) When $\mathrm{Re}\,(z)=1$, $T_z(f)\in\Lambda_{3/2}(\tilde{M})$ and $\|T_z f\|_{\Lambda_{3/2}(\tilde{M})}\leqslant A(1+|z|)^2\|f\|_\infty$.

(To prove (ii) we note that the operator T_z where $\mathrm{Re}\, z=1$ has a kernel which is like one of type 1. To prove (iii) we use the fact that since $r=2$, X_j, $j=0, 1, ..., n$, and $[X_j, X_k]$, $0<j<k\leqslant n$, span all the tangent directions. Thus to prove $T_z(f)\in\Lambda_{3/2}(\tilde{M})$, for instance, we are reduced to a case similar to a kernel of type 1.)

Now a known complex convexity argument (see Taibleson [23] or Calderón [5]) shows that $T_0(f)\in\Lambda_1(\tilde{M})$ and $\|T_0(f)\|_{\Lambda_1(\tilde{M})}\leqslant A\|f\|_\infty$. This proves the lemma.

With Lemmas 18.2 and 18.3 one shows as in § 16 that if T is of type λ with $\lambda\leqslant 2$, then T maps $L^p_\alpha(\tilde{M})$ to $L^p_{\alpha+(\lambda/r)}(\tilde{M})$, $\Lambda_\alpha(\tilde{M})$ to $\Lambda_{\alpha+(\lambda/r)}(\tilde{M})$, if $\alpha>0$, and $L^\infty(\tilde{M})$ to $\Lambda_{\lambda/r}(\tilde{M})$.

To pass to the case of general λ, one uses (18.1) to reduce the case $\lambda > 2$ to the cases $\lambda - 1$ and $\lambda - 2$, and thus obtain the result for all $\lambda > 0$. We shall describe the rest of the argument in terms of an example. Suppose that T is of type λ and $f \in L^p_\alpha(\tilde{M})$. We want to show that $Tf \in L^p_{\alpha+(\lambda/r)}(\tilde{M})$. Then it suffices to deal with $aT(f)$, where $a \in C_0^\infty(\tilde{M})$. By (18.1),

$$aT = \sum_{j=0}^{n} PT_j - S'T.$$

Here $T_j = \tilde{X}_j^2 T$, if $j = 1, ..., n$; and $T_0 = X_0 T$. In any case T_j is of type $\lambda - 2$. T is of type λ and therefore also of type $\lambda - 1$. Since P is of type 2 and S is of type 1, we are reduced to showing that T_j maps Λ_α to $\Lambda_{\alpha+((\lambda-2)/r)}$ and T maps Λ_α to $\Lambda_{\alpha+((\lambda-1)/r)}$. This reduces the problem to the case of operators of types $\lambda - 2$ and $\lambda - 1$, and the induction is complete.

Step 5. The final step in the proof of Theorem 18 requires that we observe that the properties of the mappings E and R, linking the function spaces on M with those on $\tilde{M}$, expressed in Propositions 17.3 and 17.4 go through without change.

§ 19. Estimates for $\Box_b$

Suppose M is a "partially complex" manifold of dimension $2l+1$, $l \geq 1$. (See Folland–Kohn [7], pp. 93–104.) M is then a C^∞ manifold together with a smooth sub-bundle $T_{1,0}$ (of "holomorphic" vectors) of the complex tangent bundle $\mathbf{C}T(M)$, so that $\dim_{\mathbf{C}} (T_{1,0})_\xi = l$, for $\xi \in M$; also $(T_{1,0}) \cap \overline{(T_{1,0})}_\xi = \{0\}$, and $T_{1,0}$ is integrable in the sense that the bracket of two holomorphic vector fields (cross-sections of $T_{1,0}$) is again a holomorphic vector field.

Assume we are given a hermitian metric on M, i.e. a positive definite hermitian form on the complex tangent spaces $\mathbf{C}T_\xi$, $\xi \in M$, varying smoothly with ξ, and having the property that $(T_{1,0})_\xi$ and $\overline{(T_{1,0})}_\xi$ are orthogonal.

Since our considerations are purely local (i.e. we may assume that M has, if necessary, been shrunk to a sufficiently small neighborhood of a given point ξ_0) we may construct a vector field N which is purely imaginary and is orthogonal to the spaces $T_{1,0}$ and $\overline{T}_{1,0}$ at each point. It follows that $T_{1,0}$, $\overline{T}_{1,0}$, and $\mathbf{C}N$ span the tangent space at each point.(1)

Now if Z and W are any two holomorphic vector fields, let $\varphi_\xi(Z_\xi, W_\xi)$ be the smooth function so that

$$[Z, \overline{W}]_\xi = -2\varphi_\xi(Z_\xi, W_\xi)N_\xi \quad \text{modulo} \quad (Z_1, ..., Z_l, \bar{Z}_1, ..., \bar{Z}_l). \tag{19.1}$$

Then $\varphi_\xi(,)$, $\xi \in M$, defines a hermitian form called the *Levi form*. Our basic assumption, ((19.10) below) will be in terms of the number of positive and negative eigenvalues of the

(1) N is iT in the notation of Folland-Stein [8].

Levi form. We shall choose $Z_1, Z_2, ..., Z_l$ to be an orthonormal basis of smooth holomorphic vector fields at each point. Then we have a corresponding matrix $\varphi_{jk}(\xi)$ given by

$$[Z_j, \bar{Z}_k] = -2\varphi_{jk}(\xi)N \quad \text{modulo} \quad (Z_1, ..., Z_l, \bar{Z}_1, ..., \bar{Z}_l). \tag{19.2}$$

We shall be working in a neighborhood of the fixed point ξ_0; we shall therefore make the further choice as far as the basis $Z_1, Z_2, ..., Z_l$ is concerned that at ξ_0, $\varphi_{jk}(\xi_0)$ is *diagonal*. (Note however that in general it is *not* possible to choose smooth vector fields $Z_1, Z_2, ..., Z_l$ in such a way that $\varphi_{jk}(\xi)$ is diagonal for all ξ in a neighborhood of ξ_0.)

The partially complex structure described above allows one to define the $\bar{\partial}_b$ complex; using the given metric, we can also define the dual $\mathfrak{d}_b$ complex and the resulting Laplacian. With our choice of basis we now set out to compute $\Box_b$ rather explicitly. We follow the formalism of [8], § 5 and § 13.

Let $\omega^1, ..., \omega^l$ denote a dual basis to $Z_1, ..., Z_l$. A q-form f is given by $f = \sum_{|J|=q} f_J \bar{\omega}^J$, where $J = (s_1, s_2, ..., s_q)$, with $s_1 < s_2 < ... < s_q$, $\bar{\omega}^J = \bar{\omega}^{s_1} \wedge \bar{\omega}^{s_2} ... \wedge \bar{\omega}^{s_q}$. Suppose $f = \sum_J f_J \bar{\omega}^J$, $g = \sum_J f_J \bar{\omega}^J$ are a pair of q-forms whose coefficients (the f_J and g_J) belong to $C_0^\infty(M)$. Then their inner product is given by $(f, g) = \sum_J (f_J, g_J) = \sum_J \int_M f_J(\xi)\overline{g_J(\xi)}\,d\xi$, with $d\xi$ the measure induced by the metric. When f is a function (i.e. a 0-form) the Cauchy-Riemann operator $\bar{\partial}_b$ is defined by $\bar{\partial}_b f = \sum_{j=1}^l \bar{Z}_j(f)\bar{\omega}^j$. More generally, if $f = \sum_J f_J \bar{\omega}_J$, then $\bar{\partial}_b(\sum_J f_J \bar{\omega}^J) = \sum_J (\bar{\partial}_b f) \wedge \bar{\omega}^J + \sum_J f_J \wedge \bar{\partial}_b(\bar{\omega}^J) = \sum_J \sum_{j=1}^l \bar{Z}_j(f)_J \bar{\omega}_j \wedge \bar{\omega}^J + \mathcal{E}(f)$. Here $\mathcal{E}(f)$ indicates an expression which depends linearly on f but not on derivatives of f. Similarly, the formal adjoint $\mathfrak{D}_b$ satisfies

$$\mathfrak{d}_b(\sum_J f_J \bar{\omega}^J) = -\sum_J \sum_{j=1}^l Z_j(f_J)\,\bar{\omega}^j \,\lrcorner\, \bar{\omega}^J + \mathcal{E}(f).$$

Thus

$$\mathfrak{d}_b \bar{\partial}_b(f) = -\sum_{j,k,J} Z_k \bar{Z}_j f_J \,\bar{\omega}_k \,\lrcorner\, (\bar{\omega}_j \wedge \bar{\omega}^J) + \mathcal{E}(Zf, \bar{Z}f),$$

where $\mathcal{E}(Zf, \bar{Z}f)$ represents an expression that depends on f and its first order Z and $\bar{Z}$ derivatives, but not on higher derivatives or on N derivatives of F. Similarly,

$$\bar{\partial}_b \mathfrak{d}_b(f) = -\sum_{j,k,J} \bar{Z}_j Z_k f_J \,\bar{\omega}_j \,\lrcorner\, (\bar{\omega}_k \wedge \bar{\omega}^J) + \mathcal{E}(Zf, \bar{Z}f).$$

Now if $j = k$, then

$$\bar{\omega}^k \,\lrcorner\, (\bar{\omega}^j \wedge \bar{\omega}^J) = \bar{\omega}^J, \quad \text{if } j \notin J, \quad = 0 \quad \text{if } j \in J,$$

while

$$\bar{\omega}^j \wedge (\bar{\omega}^k \,\lrcorner\, \bar{\omega}^J) = \bar{\omega}^J, \quad \text{if } j \in J, \quad = 0 \quad \text{if } j \notin J.$$

However, if $j \neq k$,

$$\bar{\omega}^k \,\lrcorner\, (\bar{\omega}^j \wedge \bar{\omega}^J) = -\bar{\omega}^j \wedge (\bar{\omega}^k \,\lrcorner\, \bar{\omega}^J).$$

Thus putting $\Box_b = \bar{\partial}_b\, \mathfrak{d}_b + \mathfrak{d}_b \bar{\partial}_b$ we get

$$\Box_b f = \Box_b^{(1)} f + \Box_b^{(2)} f + \mathcal{E}(Zf, \bar{Z}f). \tag{19.3}$$

Here $\Box_b^{(1)}$ is a "diagonal" operator (i.e. it does not mix components) given by

$$\Box_b^{(1)}(\sum_J f_J \bar{\omega}^J) = \sum_J (-\tfrac{1}{2})\left(\sum_{j=1}^{l} Z_j \bar{Z}_j + \bar{Z}_j Z_j + \varepsilon_j^J [Z_j, \bar{Z}_j]\right) f_J \bar{\omega}^J \tag{19.4}$$

with $\varepsilon_j^J = 1$ if $j \in J$, and $\varepsilon_j^J = -1$ if $j \notin J$. $\Box_b^{(2)}$ is the "non-diagonal" part given by

$$\Box_b^{(2)}(\sum_J f_J \bar{\omega}^J) = -\sum_J \sum_{j \neq k} [Z_j, \bar{Z}_k] f_J \bar{\omega}_j \wedge (\bar{\omega}_k \lrcorner \bar{\omega}^J). \tag{19.5}$$

Since the Levi form is diagonal at ξ_0, $\varphi_{jk}(\xi_0) = 0$ if $j \neq k$. However, if φ is not identically zero at ξ_0, there exists some j for which $\varphi_{jj}(\xi_0) \neq 0$. Without loss of generality we may take $j = 1$. Then by (19.2) we have

$$[Z_j, \bar{Z}_k] = d_{jk}(\xi)[Z_1, \bar{Z}_1] + \mathcal{E}(Z, \bar{Z}),$$

where $d_{jk}(\xi_0) = 0$, and $d_{jk}(\xi)$ is smooth in a neighborhood of ξ_0. Substituting this into (19.5) gives

$$\Box_b^{(2)}(\sum_J f_J \bar{\omega}^J) = -\sum_J (\sum_{j \neq k} d_{jk}(\xi)) [Z_1, \bar{Z}_1] f_J \bar{\omega}^j \wedge (\bar{\omega}^k \lrcorner \bar{\omega}^J) + \mathcal{E}(Zf, \bar{Z}f). \tag{19.5'}$$

We now introduce the real and imaginary parts of the Z_j, $j = 1, ..., l$. Let $X_1, X_2, ..., X_n$, $n = 2l$ be defined by

$$\begin{aligned} Z_j &= \tfrac{1}{2}(X_j - iX_{l+j}), \\ \bar{Z}_j &= \tfrac{1}{2}(X_j + iX_{l+j}), \quad j = 1, 2, ..., l. \end{aligned} \tag{19.6}$$

The passage from the complex vector fields to the real vector fields induces a passage from $l \times l$ hermitian matrices to $2l \times 2l$ real skew symmetric matrices given as follows. If a is any $l \times l$ hermitian matrix the corresponding (real skew-symmetric) $2l \times 2l$ matrix $a^\dagger$ is given by

$$a^\dagger = \left(\begin{array}{c|c} \mathrm{Im}(a) & \mathrm{Re}(a) \\ \hline -\mathrm{Re}(a) & \mathrm{Im}(a) \end{array}\right). \tag{19.7}$$

From (19.6) it follows directly that $Z_j \bar{Z}_j + \bar{Z}_j Z_j = \tfrac{1}{2}(X_j^2 + X_{j+l}^2)$ $[Z_j, \bar{Z}_j] = \tfrac{i}{2}[X_j, X_{j+l}]$. Thus (19.4) becomes

$$\Box_b^{(1)}(\sum_J f_J \bar{\omega}^J) = -\tfrac{1}{4} \sum_J \mathcal{L}_J(f_J)\, \bar{\omega}^J, \tag{19.8}$$

with

$$\mathcal{L}_J = \sum_{j=1}^{n} X_j^2 + \frac{i}{2} \sum_{j,k} \gamma_{j,k}^J [X_j, X_k],$$

where

$$\gamma^J = (\varepsilon^J)^\dagger = \left(\begin{array}{c|c} 0 & \varepsilon^J \\ \hline -\varepsilon^J & 0 \end{array}\right),$$

and ε^J is the diagonal matrix $\varepsilon_{jk}^J = \varepsilon_j^J \delta_{jk}$, $1 \leqslant j, k \leqslant l$.

The identity (19.2) becomes

$$[X_j, X_k] = 4i(\varphi_{jk}(\xi))^\dagger N \quad \text{modulo} \quad (X_1, X_2, \dots, X_n). \tag{19.9}$$

We now state the main hypothesis on the Levi form φ for a given q at each point $\xi \in M$:

$$p_1 \geqslant \max(q+1, l+1-q), \quad \text{or} \quad p_2 \geqslant \min(q+1, l+1-q), \tag{19.10}$$

where p_1 is the larger of the number of eigenvalues of the same sign, and p_2 is the number of pairs of eigenvalues of opposite sign.

(19.11) LEMMA. *If the Levi form φ at ξ_0 satisfies the hypothesis* (19.10), *then*

$$|\operatorname{tr}(\gamma^J \cdot \varphi^\dagger)| \leqslant \sigma \|\varphi^\dagger\|_1, \tag{19.12}$$

for some σ, $0 \leqslant \sigma < 1$.

Proof. We obviously have $\operatorname{tr}(\gamma^J\varphi^\dagger) = 2 \operatorname{tr}(\varepsilon^J\varphi_R)$, where φ_R is the real part of the hermitian matrix φ. We have assumed that φ at ξ_0 is diagonal. Let $\lambda_1, \dots, \lambda_l$ by its eigenvalues. Then $\operatorname{tr}(\varepsilon^J\varphi_R) = \operatorname{tr}(\varepsilon^J\varphi) = \sum_{j=1}^l \varepsilon_j^J \lambda_j$. By checking the four cases implicit in (19.10), one can see that it is not possible for all the $\varepsilon_j^J\lambda_j$, $j = 1, 2, \dots, n$ to be positive, or all $\varepsilon_j^J\lambda_j$ to be negative. Therefore, there is a number σ, $0 \leqslant \sigma < 1$ such that $|\sum \varepsilon_j^J \lambda_j| \leqslant \sigma \sum_{j=1}^l |\lambda_j|$. However, the eigenvalues of $\varphi^\dagger$ are $\{\pm i\lambda_j\}_{j=1,\dots,l}$, so that $\|\varphi^\dagger\| = 2 \sum_{j=1}^l |\lambda_j|$, which proves the lemma.

By the lemma the linear functional $x \to \operatorname{tr}(\gamma^J x)$ has norm not exceeding σ, when defined on the one-dimensional subspace spanned by $\varphi^\dagger(\xi_0)$ in the space of $n \times n$ real skew symmetric matrices. For the present purpose we take the norm on this space of matrices to be the trace norm, $\| \, \|_1$. Thus by the Hahn-Banach theorem the linear function can be extended to the whole space of skew symmetric matrices without increasing the norm. Invoking the duality (2.2), this means that there is a real skew-symmetric $2l \times 2l$ matrix $h = h_{jk}^J$ with the properties

$$\begin{cases} \operatorname{tr}(h^J\varphi^\dagger(\xi_0)) = 0 \\ \|\gamma^J + h^J\| \leqslant \sigma < 1. \end{cases} \tag{19.13}$$

Next, choose a smooth function $h^J(\xi)$ so that $h^J(\xi_0) = h^J$ and

$$\operatorname{tr}(h^J(\xi)\varphi^\dagger(\xi)) = 0 \tag{19.13$'$}$$

for all ξ in some neighborhood of ξ_0.

Because of (19.8) and (19.3) we can write

$$-4\Box_b(f) = \mathcal{L}(f) + \mathcal{E}(Xf), \tag{19.14}$$

where $\mathcal{L}(f) = \sum X_j^2 f + \frac{1}{2} \sum c_{jk}[X_j, X_k] f$ and $\mathcal{E}(Xf)$ is an error term depending linearly on f and $X_j f$, $j = 1, 2, \ldots, n$. In fact, $c_{jk}(\xi)(\sum f_J \bar{\omega}^J) = i \sum_J (\gamma_{jk}^J(\xi) + h_{jk}^J(\xi)) f_J \bar{\omega}^J$ + terms coming from $\Box_b^{(2)}$. (Note that $\sum h_{jk}^J(\xi)[X_j, X_k]_\xi = \mathcal{E}(X_\xi)$ by (19.13′).) However, by (19.5′) the coefficients of the non-diagonal part, $\Box_b^{(2)}$, vanish at $\xi = \xi_0$, modulo an error term which may be absorbed in $\mathcal{E}(Xf)$.

Thus for each j, k the value of the function c_{jk} at ξ_0 is a *diagonal* matrix. Furthermore,

$$(c_{jk}(\xi_0))_{J,J} = i(\gamma_{jk}^J(\xi_0) + h_{jk}^J(\xi_0)).$$

Hence by (19.13) and the results of Part I and in particular, Theorem 1, the condition (15.2′) is satisfied at ξ_0, and by continuity in a neighborhood of ξ_0; we restrict ourselves to this neighborhood in what follows. (For this application observe that since the Levi form φ is non-zero, by (19.9) it follows that $X_1, X_2, \ldots, X_n$ together with their commutators of length 2 span the tangent space at each point.) We can therefore invoke Theorem 15 to obtain regularity properties of solutions of $\Box_b$.

THEOREM 19. *Assume that the condition* (19.10) *is satisfied. If f and g are q-forms each in L^p with $\Box_b f = g$, then the following hold.*

(a) *If $g \in L_\alpha^p(M)$, then $f \in L_{\alpha+1}^p(M)$.*
(b) *If $g \in \Lambda_\alpha(M)$, then $f \in \Lambda_{\alpha+1}(M)$, $\alpha > 0$.*
(c) *If $g \in L^\infty(M)$, then $f \in \Lambda_1(M)$.*
(d) *If $g \in S_k^p(M)$, then $af \in S_{k+2}^p(M)$ for all $a \in C_0^\infty(M)$.*

Proof. By Theorem 15 for each $a \in C(_0^\infty M)$ there exists a parametrix P which is smoothing of order 2 and a smoothing operator S' of order 1, so that $P\mathcal{L} = aI + S'$. Since $-4\Box_b = \mathcal{L} + \mathcal{E}(Xf)$, we take $P_1 = -4P$ and $S_1 = S' - P(\mathcal{E}(Xf))$. Then

$$P_1 \Box_b f = af + S_1(f), \tag{19.15}$$

where S_1 is then also smoothing of order 1. (See Proposition 17.12.)

Next we can iterate the identity (19.15) as in the proof of the corollary to Theorem 15 to obtain error terms which are smoothing of preassigned order. Finally the rest of the proof of the theorem follows as in the proof of Theorem 16.

The above also gives optimal regularity results for global solutions of $\bar{\partial}_b f = g$ on compact manifolds, as in [8], § 17; we shall not enter into the details here.

References

[1]. Abrahams, R. & Robbin, J., *Transversal Mappings and Flows*, W. A. Benjamin, Inc. New York, 1967.

[2]. Baouendi, M. S. & Goulaouic, G. Non-analytic hypoellipticity for some degenerate elliptic operators. *Bull. Amer. Math Soc.*, 78 (1972), 483–486.

[3]. Boutet de Monvel, L. & Trèves, F., On a class of pseudodifferential operators with double characteristics *Invent. Math.*, 24 (1974) 1–34.

[4]. Calderón, A. P., Lebesgue spaces of differentiable functions and distributions., *Proc. Symp. Pure Math.*, Vol. 4, Amer. Math. Soc. 1961, 33–49.

[5]. —— Intermediate spaces and interpolation, the complex method, *Studia Math.*, 24, (1964), 113–190.

[6]. Folland, G. B., Subelliptic estimates and function spaces on nilpotent Lie groups, *Arkiv f. Mat.*, *13*, 1975, 161–207.

[7]. Folland, G. B. & Kohn, J. J., *The Neumann Problem for the Cauchy-Riemann Complex.* Ann. of Math. Studies 75, Princeton Univ. Press, Princeton, N.J. 1972.

[8]. Folland, G. B. & Stein, E. M., Estimates for the $\bar{\partial}_b$-complex and analysis on the Heisenberg group. *Comm. Pure Appl. Math.*, 27 (1974), 429–522.

[9]. Greiner, P. & Stein, E. M., A parametrix for the $\bar{\partial}$-Neumann problem. Rencontre sur l'Analyse complexe à plusieurs variables ..., Montreal, 1975, 49–63.

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

[11]. Jacobson, N., *Lie algebras.* Interscience Tracts in Pure and Applied Mathematics, Vol. 10, John Wiley and Sons, New York 1962.

[12]. Knapp, A. & Stein, E. M., Intertwining operators for semisimple groups, *Ann. of Math.*, 93 (1971), 489–578.

[13]. Kohn, J. J., Boundaries of complex manifolds. *Proc. Conference on Complex Manifolds, Minneapolis*, 1964, 81–94.

[14]. —— Complex hypoelliptic operators, *Symposia Mathematica*, 7 (1971), 459–468.

[15]. —— Pseudo-differential operators and hypoellipticity *Proc. Symp. Pure Math.*, 23, Amer. Math. Soc., 1973, 61–69.

[16]. Korányi, A. & Vagi, S., Singular integrals in homogeneous spaces and some problems of classical analysis. *Ann. Scuola Norm. Sup.* Pisa, 25 (1971), 575–648.

[17]. Lions, J. L. & Peetre, J., Sur une classe d'espaces d'interpolation, *Publ. Math Inst. Hautes Etudes Sci. Paris*, 19 (1964), 5–68.

[18]. O'Neil, R., Two elementary theorems of the interpolation of linear operators, *Proc. Amer. Math Soc.*, 17 (1966), 76–82.

[19]. Pukanszky, L., *Lecons sur les Representations des Groupes.* Monographies de La Societé Mathématique de France, Dunod, Paris, 1967.

[20]. Radkevich, E. V., Hypoelliptic operators with multiple characteristics, *Mat. Sbornik*, 79, (121) (1969), 193–216. (*Math. USSR Sbornik*, 8 (1969), 181–205.

[21]. Serre, J.-P., *Lie algebras and Lie groups*, Benjamin, New York, 1965.

[22]. Stein, E. M., *Singular integrals and differentiability properties of functions.* Princeton Univ. Press, Princton, 1970.

[23]. Taibleson, M. H., Translation invariant operators, duality, and interpolation, II. *J. Math. Mech.*, 14 (1965), 821–840.

Received June 23, 1975