Subelliptic estimates and function spaces on nilpotent Lie groups

G. B. Folland*

Introduction

In recent years there has been considerable activity in the study of hypoelliptic but non-elliptic partial differential equations. (We recall that a differential operator $\mathscr{L}$ on a manifold M is said to be hypoelliptic if for any open set $U \subset M$ and distributions f, g on U satisfying $\mathscr{L}f=g$ on U, $f \in \mathscr{C}^\infty(U)$ implies $g \in \mathscr{C}^\infty(U)$.) One of the major ideas in this field is that of obtaining control over the characteristic directions of a differential operator by conditions involving commutators of vector fields or pseudodifferential operators. The prototype of such results is the following theorem of Hörmander [10]:

(0.1) **Proposition.** *Let $X_0, X_1, \dots, X_n$ be real vector fields on an open set $U \subset \mathbf{R}^N$, and let $\mathscr{V}_k$ be the linear span of the vector fields $X_{i_1}, [X_{i_1}, X_{i_2}], \dots, [[\dots[X_{i_1}, X_{i_2}], \dots \dots, X_{i_{k-1}}], X_{i_k}]$ $(0 \leqq i_j \leqq n,\ 1 \leqq j \leqq k)$. Suppose there is an integer m such that $\mathscr{V}_m$ spans the tangent space at every point of U. Then the operator $\mathscr{L} = X_0 + \sum_1^n X_j^2$ is hypoelliptic on U.*

If the hypotheses of this theorem are satisfied, the more refined regularity properties of $\mathscr{L}$ (in terms of L^2 estimates, say) depend strongly on the integer m: roughly speaking, the larger m is, the "weaker" $\mathscr{L}$ is. We refer to Hörmander [10] for a precise interpretation of this statement.

Similar ideas occur in the study of the $\bar\partial$ Neumann problem and $\bar\partial_b$ complex in several variables (cf. Folland—Kohn [5]) and their analogues for more general differential complexes (Guillemin—Sternberg [8]), and in the work of the Russian school on hypoelliptic equations (see Oleĭnik—Radkevič [22] and the references given there). Operators of Hörmander's type are discussed from the point of view

* Research partially supported by National Science Foundation Grant GP—38220.

of potential theory in Bony [1], and some applications to infinite-dimensional group representations have recently been given by Jørgensen [12].

In the theory of *elliptic* operators the constant-coefficient operators serve as a useful class of models for the general situation: constant-coefficient operators are amenable to treatment by the techniques of Euclidean harmonic analysis (Fourier transforms, convolution operators, etc.), and the results obtained thereby can usually be extended to the variable-coefficient case by perturbation arguments. Now, a constant-coefficient operator is nothing more than a translation-invariant operator on the Abelian Lie group $\mathbf{R}^N$. From this point of view, it is natural to attempt to construct a class of models for non-elliptic operators of the sort discussed above among the translation-invariant operators on certain non-Abelian Lie groups. The Lie algebras of the groups involved should have a structure which reflects the behavior of the commutators in the original problem and the groups themselves should admit a "harmonic analysis" which will produce results similar to those of the Euclidean case. A particular case of this program, has been carried out in considerable detail in Folland—Stein [6], [7], in which sharp L^p and Lipschitz (or Hölder) estimates for the $\bar{\partial}_b$ complex on the boundary of a complex domain with nondegenerate Levi form are obtained by using certain left-invariant operators on the Heisenberg group as models.

The purpose of this paper is to construct a general theory of "subelliptic" regularity on a class of Lie groups which should be sufficiently broad to admit a wide variety of applications to more general problems, namely the class of "stratified groups". We call a Lie group G *stratified* if it is nilpotent and simply connected and its Lie algebra $\mathfrak{g}$ admits a vector space decomposition $\mathfrak{g}=V_1\oplus\ldots\oplus V_m$ such that $[V_1, V_k]=V_{k+1}$ for $1\leqq k<m$ and $[V_1, V_m]=\{0\}$. On such groups there is a natural notion of homogeneity which enables one to duplicate many of the standard constructions of Euclidean space (for example, a theory of singular integral operators parallel to the Calderón—Zygmund theory). Also, if we choose a basis $X_1, \ldots, X_n$ for V_1, the operator $\mathcal{J}=-\sum_1^n X_j^2$ (which is hypoelliptic by Hörmander's theorem) turns out to play much the same fundamental role on G as (minus) the ordinary Laplacian $-\sum_1^N(\partial/\partial x_j)^2$ does on $\mathbf{R}^N$. We call $\mathcal{J}$ a *sub-Laplacian* for G.

The plan of the paper is as follows. In Section 1 we present the necessary background material concerning homogeneous structures on nilpotent Lie groups. Much of this is not new, but we include most of the proofs in the interest of making the exposition reasonably self-contained. In Section 2 we prove that homogeneous hypoelliptic operators on nilpotent groups have homogeneous fundamental solutions, and we give some examples. The main theme of the paper begins its development in Section 3, where we consider the diffusion semigroup generated by the sub-Laplacian $\mathcal{J}$ on a stratified group and use it to define complex powers of $\mathcal{J}$ in accordance with the general theory of fractional powers of operators due to

Komatsu and others. In section 4 we define analogues of the classical L^p Sobolev or potential spaces in terms of fractional powers of $\mathcal{J}$ and extend several basic theorems from the Euclidean theory of differentiability to these spaces: interpolation properties, boundedness of singular integrals, representations in terms of derivatives, localizability, and imbedding theorems. We also relate our new Sobolev spaces to the classical ones. In Section 5 we define spaces of functions satisfying certain Lipschitz (or Hölder) conditions which are compatible with the homogeneous structure on the group. We prove boundedness theorems for homogeneous integral operators in terms of these spaces, relate L^p conditions to Lipschitz conditions by an extension of the Sobolev imbedding theorem, and compare our new Lipschitz spaces to the classical ones. Finally, in Section 6 we apply the preceding material to derive sharp L^p and Lipschitz estimates for homogeneous hypoelliptic operators on stratified groups.

Some of our results have been obtained independently by R. S. Strichartz (personal communication).

This paper and its author both owe a great deal to Elias M. Stein. Most of the basic ideas herein were developed through conversations with him (indeed, the germ of these ideas was already present in his lecture at the Nice congress [26]), and many of the results and techniques are extensions of those in our joint work [7]. I also wish to thank Robert T. Moore for several helpful conversations.

1. Homogeneous structures on nilpotent groups

Let $\mathfrak{g}$ be a real finite-dimensional Lie algebra. A *family of dilations* on $\mathfrak{g}$ is a one-parameter family $\{\gamma_r: 0<r<\infty\}$ of automorphisms of $\mathfrak{g}$ of the form $\gamma_r=\exp(A \log r)$ where A is a diagonalizable linear transformation of $\mathfrak{g}$ with positive eigenvalues. If $\{\gamma_r\}$ is a family of dilations, then so is $\{\tilde{\gamma}_r\}$ where $\tilde{\gamma}_r=\gamma_{r^\alpha}=\exp(\alpha A \log r)$ for any $\alpha>0$. Hence, by adjusting α if necessary, we always assume that the smallest eigenvalue of A is 1.

It is easy to see that if $\mathfrak{g}$ has a family of dilations then $\mathfrak{g}$ is nilpotent. Otherwise, one could find arbitrarily long sequences $X_1, \ldots, X_m \in \mathfrak{g}$ such that X_j is an eigenvector of A with eigenvalue α_j and $Y=[[\ldots[X_1, X_2], \ldots, X_{m-1}], X_m]\neq 0$. Since γ_r is an automorphism, $\gamma_r Y=r^{\Sigma\alpha_j}Y$, so Y is also an eigenvector of A with eigenvalue $\sum_1^m \alpha_j \geqq m$. But this is possible only for finitely many m. On the other hand, it is known (J. Dyer [4]) that not every nilpotent Lie algebra admits dilations.

Let $\mathfrak{g}$ be a nilpotent Lie algebra with dilations $\{\gamma_r\}$, and let G be the corresponding simply connected Lie group. Since $\mathfrak{g}$ is nilpotent, the exponential map $\exp: \mathfrak{g}\to G$ is a diffeomorphism (cf. Hochschild [9]), and the dilations γ_r lift via exp to give a one-parameter group of automorphisms of G, which we still denote by γ_r. We fix once and for all a (bi-invariant) Haar measure dx on G (which is just the lift of Lebesgue measure on $\mathfrak{g}$ via exp). The number $Q=\text{trace}\,(A)$ will be called the *homogeneous dimension* of G (with respect to the dilations $\{\gamma_r\}$). The reason for this is that

$$d(\gamma_r x) = r^Q\, dx. \tag{1.1}$$

We note that the homogeneous dimension is generally greater than (and never less than) the Euclidean dimension.

Some matters of notation: We shall denote the identity element of G by 0, even though we write the group law multiplicatively, in order to emphasize the similarity to Euclidean space. We use the standard notations $\mathcal{D}$, $\mathcal{E}$, $\mathcal{D}'$, and $\mathcal{E}'$ for the spaces of $\mathcal{C}^\infty$ functions with compact support, $\mathcal{C}^\infty$ functions, distributions, and distributions with compact support on G, with the usual locally convex topologies (cf. Schwartz [23]). (However, we often write $\mathcal{C}^\infty$ instead of $\mathcal{E}$.) The pairing of $\tau\in\mathcal{D}'$ with $u\in\mathcal{D}$ will be denoted $\langle\tau, u\rangle$; we shall use duality over $\mathbf{R}$ throughout, so that this pairing is bilinear rather than sesquilinear. δ will denote the Dirac distribution: $\langle\delta, u\rangle=u(0)$ for $u\in\mathcal{D}$. Also, $\mathcal{C}_0$ will denote the space of continuous functions on G vanishing at infinity, with the uniform topology, and L^p $(1\leqq p\leqq\infty)$ will denote the standard L^p space with respect to Haar measure. Finally, we identify the Lie algebra $\mathfrak{g}$ with the *left-invariant* vector fields on G.

A measurable function f on G will be called *homogeneous of degree* λ $(\lambda\in\mathbf{C})$ if $f\circ\gamma_r=r^\lambda f$ for all $r>0$. Likewise, a distribution $\tau\in\mathcal{D}'$ will be called *homogeneous of degree* λ if $\langle\tau, u\circ\gamma_r\rangle=r^{-Q-\lambda}\langle\tau, u\rangle$ for all $u\in\mathcal{D}$ and $r>0$. In view of (1.1), these definitions are consistent. A distribution which is $\mathcal{C}^\infty$ away from 0 and homogeneous of degree $\alpha-Q$ will be called a *kernel of type* α.

A differential operator D will be called *homogeneous of degree* λ if $D(u\circ\gamma_r)=r^\lambda(Du)\circ\gamma_r$ for all $u\in\mathcal{D}$, $r>0$. In particular, $X\in\mathfrak{g}$ is homogeneous of degree λ if and only if X is an eigenvector of A with eigenvalue λ. It is then clear that if K is a kernel of type α and D is homogeneous of degree λ, then DK is a kernel of type $\alpha-\lambda$.

Let $\|\cdot\|$ denote a Euclidean norm on $\mathfrak{g}$ with respect to which the eigenspaces of A are mutually orthogonal; we may also regard $\|\cdot\|$ as a function on G. In addition, we shall need a "norm" on G which respects the homogeneous structure. Namely, we define a *homogeneous norm* to be a continuous function from G to $[0, \infty)$ which is $\mathcal{C}^\infty$ away from 0 and homogeneous of degree 1, and which satisfies (a) $|x|=0$ if and only if $x=0$, (b) $|x|=|x^{-1}|$ for all x. Homogeneous norms always exist. Indeed, any $X\in\mathfrak{g}$ can be written as $X=X_1+\ldots+X_m$ where each X_j is an eigenvector of A with eigenvalue α_j, and since the X_j's are mutually orthogonal, $\|\gamma_r X\|=(\sum r^{2\alpha_j}\|X_j\|^2)^{1/2}$ is a strictly increasing function of r. Since moreover $\gamma_r\circ\exp=\exp\circ\gamma_r$, we may define $|x|$ for $x\neq 0\in G$ to be the unique $r>0$ such that $\|\gamma_r^{-1}x\|=1$. We assume henceforth that we have fixed a homogeneous norm on G.

(1.2) **Lemma.** *$\{x\in G: |x|\leqq 1\}$ is compact.*

Proof. Since $\{x: \|x\|=1\}$ is compact and disjoint from 0, $|\cdot|$ assumes a positive minimum M on it. Since $\{x: |x|\leqq M\}$ is connected (each such x is connected to 0 by the arc $\{\gamma_r x: 0\leqq r\leqq 1\}$), $\{x: |x|\leqq M\}\subset\{x: \|x\|\leqq 1\}$. But then $\{x: |x|\leqq 1\}\subset \subset\{x: \|\gamma_M x\|\leqq 1\}$, so $\{x: |x|\leqq 1\}$ is closed and bounded, hence compact.

(1.3) **Lemma.** *Let a be the largest eigenvalue of A. There exist $C_1, C_2>0$ such that $C_1\|x\|\leqq|x|\leqq C_2\|x\|^{1/a}$ whenever $|x|\leqq 1$.*

Proof. As above, we can write $x\in G$ as $x=\exp(X_1+\ldots+X_m)$ where X_j is an eigenvector of A with eigenvalue a_j, $1\leqq a_j\leqq a$, and $\|\gamma_r x\|=(\sum r^{2a_j}\|X_j\|^2)^{1/2}$. Hence for $r\leqq 1$ we have $r^a\|x\|\leqq\|\gamma_r x\|\leqq r\|x\|$. By Lemma (1.2), $\|\cdot\|$ assumes a positive maximum C_1^{-1} and a positive minimum C_2^{-a} on $\{x: |x|=1\}$. Any x with $|x|\leqq 1$ can be written as $x=\gamma_{|x|}y$ with $|y|=1$, so that

$$\|x\| \leqq |x|\,\|y\| \leqq C_1^{-1}|x|, \quad \|x\| \geqq |x|^a\|y\| \geqq C_2^{-a}|x|^a.$$

(1.4) **Proposition.** *There is a constant $C>0$ such that $|xy|\leqq C(|x|+|y|)$ for all $x, y\in G$.*

Proof. By Lemma (1.2), the set $\{(x, y)\in G\times G: |x|+|y|=1\}$ is compact, so the function $(x, y)\to|xy|$ assumes a finite maximum C on it. Then, given any $x, y\in G$, set $r=|x|+|y|$. It follows that

$$|xy| = r|\gamma_r^{-1}(xy)| = r|(\gamma_r^{-1}x)(\gamma_r^{-1}y)| \leqq Cr = C(|x|+|y|).$$

We now prove a number of facts about homogeneous functions and distributions.

(1.5) **Proposition** (Knapp—Stein [13]). *Let f be a homogeneous function of degree $-Q$ which is locally integrable away from 0. There exists a constant $M(f)$, the "mean value" of f, such that*

$$\int f(x)g(|x|)\,dx = M(f)\int_0^\infty r^{-1}g(r)\,dr$$

for all functions g on $(0, \infty)$ such that either side makes sense.

Proof. Define $L:(0, \infty)\to \mathbf{C}$ by

$$L(r) = \int_{1\leqq|x|\leqq r} f(x)\,dx \quad \text{if} \quad r\geqq 1$$

$$-\int_{r\leqq|x|\leqq 1} f(x)\,dx \quad \text{if} \quad r<1.$$

By using (1.1) one easily checks that $L(rs)=L(r)+L(s)$ for all $r, s>0$. Since L is continuous, it follows that $L(r)=L(e)\log r$. We take $M(f)=L(e)$; the assertion is then clear when g is the characteristic function of an interval, and it follows in general by taking linear combinations and limits.

(1.6) **Corollary.** *Let* $C_0 = M(|\cdot|^{-Q})$. *Then if* $\alpha \in \mathbf{C}$ *and* $0 < a < b < \infty$,

$$\int_{a \leq |x| \leq b} |x|^{\alpha - Q}\, dx = \begin{cases} C_0 \alpha^{-1}(b^\alpha - a^\alpha) & \text{if } \alpha \neq 0, \\ C_0 \log(b/a) & \text{if } a = 0. \end{cases}$$

Proof. Take $g(r) = r^a$ times the characteristic function of $[a, b]$.

(1.7) **Corollary.** *Suppose f is a measurable function such that* $|f(x)| = O(|x|^{\alpha - Q})$. *If* $\alpha > 0$ *then f is locally integrable at* 0, *and if* $\alpha < 0$ *then f is locally integrable at infinity.*

If $f \in \mathscr{C}^\infty(G - \{0\})$ is homogeneous of degree $-Q$, then f is not a distribution since it is not locally integrable at 0. However, if $M(f) = 0$ there is canonically associated to f a kernel of type 0, denoted $PV(f)$, which is defined by

$$\langle PV(f), u \rangle = \lim_{\varepsilon \to 0} \int_{|x| > \varepsilon} f(x) u(x)\, dx \quad (u \in \mathscr{D}).$$

To see that this is well defined, we note that

$$\begin{aligned} \langle PV(f), u \rangle &= \lim_{\varepsilon \to 0} \int_{\varepsilon < |x| \leq 1} f(x)[u(x) - u(0)]\, dx + \int_{|x| \geq 1} f(x) u(x)\, dx \\ &= \int_{|x| < 1} f(x)[u(x) - u(0)]\, dx + \int_{|x| \geq 1} f(x) u(x)\, dx. \end{aligned}$$

The last integrals are absolutely convergent by Corollary (1.7), since $u(x) - u(0) = O(\|x\|) = O(|x|)$ by Lemma (1.3). It is easy to check that $PV(f)$ is homogeneous of degree $-Q$.

(1.8) **Proposition** (Folland—Stein [7]). *Let K be a kernel of type α which agrees with* $f \in \mathscr{C}^\infty(G - \{0\})$ *away from* 0. *Then*

(a) *f is homogeneous of degree* $\alpha - Q$.

(b) *If* $\operatorname{Re} \alpha > 0$, *then* $f \in L^1(\text{loc})$ *and* $K = f$.

(c) *If* $\alpha = 0$, *then* $M(f) = 0$ *and* $K = PV(f) + C\delta$ *for some constant C.*

Proof. (a) is obvious. For (b) we observe that f is locally integrable by Corollary (1.7), so $f \in \mathscr{D}'$. $K - f$ is thus a distribution supported at 0, that is, a linear combination of δ and its derivatives (cf. [23]). Every derivative at 0 is a sum of homogeneous terms of positive degree, and hence for any $u \in \mathscr{D}$,

$$\langle K - f, u \circ \gamma_r \rangle = O(1) \quad \text{as} \quad r \to \infty.$$

But $K-f$ is homogeneous of degree $\alpha-Q$, so

$$\langle K-f, u\circ\gamma_r\rangle = r^\alpha\langle K-f, u\rangle,$$

which is a contradiction for $\operatorname{Re}\alpha>0$ unless $K-f=0$.

As for (c), if $M(f)=0$ then $K-PV(f)$ is homogeneous of degree $-Q$ and supported at 0, hence is a multiple of δ. To show $M(f)=0$, consider another distribution F which agrees with f away from 0, namely

$$\langle F, u\rangle = \int_{|x|<1} f(x)[u(x)-u(0)]\,dx + \int_{|x|\geqq 1} f(x)u(x)\,dx \quad (u\in\mathscr{D}),$$

which is well defined by the reasoning preceding this proposition. Hence, as above, $K-F$ is a linear combination of δ and its derivatives, so since K is homogeneous of degree $-Q$,

$$\langle F, u\circ\gamma_r\rangle-\langle F, u\rangle = \langle F-K, u\circ\gamma_r\rangle-\langle F-K, u\rangle = O(1) \quad \text{as} \quad r\to\infty.$$

But in fact

$$\langle F, u\circ\gamma_r\rangle-\langle F, u\rangle = -u(0)\int_{(1/r)\leqq|x|\leqq 1} f(x)\,dx = u(0)M(f)\log r,$$

which is a contradiction when $u(0)\neq 0$ unless $M(f)=0$.

Kernels of type 0 are thus a natural generalization of the classical Calderón—Zygmund singular integral kernels (cf. Stein [25]), and we have the following L^p boundedness theorem for the convolution operators defined by them.

(1.9) **Proposition.** *Let $K=PV(f)+C\delta$ be a kernel of type* 0. *The mapping $T\colon u\to u*K$ $(u\in\mathscr{D})$ extends to a bounded operator on L^p, $1<p<\infty$. In fact, set $T_\varepsilon u= =u*f_\varepsilon+Cu$ where $f_\varepsilon(x)=f(x)$ for $|x|>\varepsilon$ and $=0$ for $|x|\leqq\varepsilon$. Then T_ε is bounded on L^p uniformly in ε, and T is the strong L^p limit of T_ε as $\varepsilon\to 0$. Likewise for the mapping $\tilde{T}\colon u\to K*u$.*

Proof. See Knapp—Stein [13] for the case $p=2$, and Coifman—Weiss [3] or Korányi—Vági [21] for the extension to other values of p.

There is a corresponding result for kernels of type $\alpha>0$. We deduce it as a corollary of the following generalization of Young's inequality, which is implicitly stated in Stein [25] and Folland—Stein [7]. We recall that if f is a measurable function on G, its *distribution function* $\beta_f\colon (0,\infty)\to[0,\infty]$ is defined by $\beta_f(a)=|\{x\colon |f(x)|>a\}|$, where $|E|$ is the measure of E. f is said to be in *weak* L^r $(1\leqq r<\infty)$ if for some $C>0$, $\beta_f(a)\leqq(C/a)^r$ for all $a>0$. The smallest such C will be denoted by $[f]_r$. $[\ ]_r$ does not satisfy the triangle inequality; however, it defines a topology on weak L^r which coincides with a Banach space topology in case $r>1$. Moreover, if $f\in L^r$ then $f\in$ weak L^r and $[f]_r\leqq\|f\|_r$, and we have

$$\|f\|_r = \left(r\int_0^\infty a^{r-1}\beta_f(a)\,da\right)^{1/r}.$$

For proofs of these facts, see Stein—Weiss [28].

(1.10) **Proposition.** *Suppose* $1\leqq p\leqq\infty$, $1<r<\infty$, *and* $q^{-1}=p^{-1}+r^{-1}-1>0$. *If* $f\in L^p$ *and* $g\in$ *weak* L^r *then* $f*g$ *exists a.e. and is in weak* L^q, *and there exists* $C_1=C_1(p,r)>0$ *such that* $[f*g]_q\leqq C_1\|f\|_p[g]_r$. *Moreover, if* $p>1$ *then* $f*g\in L^q$, *and there exists* $C_2=C_2(p,r)>0$ *such that* $\|f*g\|_q\leqq C_2\|f\|_p[g]_r$. *The same results hold with* $f*g$ *replaced by* $g*f$.

Proof. We first observe that the strong result for $p>1$ follows from the weak result by the Marcinkiewicz interpolation theorem with explicit bounds (cf. Zygmund [31]). Suppose then that $f\in L^p$ and $g\in$weak L^r; we may assume $\|f\|_p=1$ and $[g]_r=1$. Given $a>0$, set $M=(a/2)^{q/r}(r/q)^{q/rp'}$ where $p'=p/(p-1)$ is the conjugate exponent to p. Define $g_1(x)$ to be $g(x)$ if $|g(x)|>M$ and 0 otherwise, and set $g_2=g-g_1$. Since

$$\beta_{f*g}(a)\leqq\beta_{f*g_1}(a/2)+\beta_{f*g_2}(a/2),$$

it suffices to estimate each term on the right. By Hölder's inequality, $|f*g_2(x)|\leqq\leqq\|f\|_p\|g_2\|_{p'}=\|g_2\|_{p'}$. However, since $r^{-1}-(p')^{-1}=q^{-1}>0$, we have $p'-r>0$, hence

$$\|g_2\|_{p'}^{p'}=p'\int_0^M\alpha^{p'-1}\beta_g(\alpha)\,d\alpha\leqq p'\int_0^M\alpha^{p'-1-r}\,d\alpha$$

$$=\frac{p'}{p'-r}M^{p'-r}=\frac{q}{r}M^{rp'/q}=(a/2)^{p'}.$$

Thus $f*g_2(x)$ exists for every x and $|f*g_2(x)|\leqq a/2$, so $\beta_{f*g_2}(a/2)=0$. On the other hand, since $r>1$,

$$\|g_1\|_1=\int_M^\infty\beta_g(\alpha)\,d\alpha\leqq\int_M^\infty a^{-r}\,da=(1-r)^{-1}M^{1-r}.$$

Thus by Young's inequality, $f*g_1$ exists a.e. and is in L^p, and $\|f*g_1\|_p\leqq\|f\|_p\|g_1\|_1\leqq\leqq(1-r)^{-1}M^{1-r}$. But then

$$\beta_{f*g_1}(a/2)\leqq\left(\frac{2\|f*g_1\|_p}{a}\right)^p\leqq\left(\frac{2}{a}\right)^p\left(\frac{1}{1-r}\right)^p\left(\frac{a}{2}\right)^{pq(1-r)/r}\left(\frac{r}{q}\right)^{pq(1-r)/rp'}\leqq C_1(p,r)a^{-q}$$

and the proof is complete. (The proof is the same for $g*f$.)

(1.11) **Proposition.** *Suppose* $0<\alpha<Q$, $1<p<Q/\alpha$, *and* $q^{-1}=p^{-1}-(\alpha/Q)$, *and let* K *be a kernel of type* α. *If* $f\in L^p$ *then* $f*K$ *and* $K*f$ *exist a.e. and are in* L^q, *and there is a constant* $C_p>0$ *such that* $\|f*K\|_q\leqq C_p\|f\|_p$ *and* $\|K*f\|_q\leqq C_p\|f\|_p$.

Proof. By Proposition (1.8), K is a function, so by Proposition (1.10) it suffices to show that $K\in$weak L^r where $r=Q/(Q-\alpha)$. But $|K(x)|$ is dominated by $|x|^{\alpha-Q}$, so $\beta_K(a)$ is dominated by the measure of $\{x:|x|<a^{1/(\alpha-Q)}\}$. By Corollary (1.6), this number is $C_0a^{Q/(\alpha-Q)}=C_0a^{-r}$.

We recall some facts about convolution of distributions (see Schwartz [23] for the case $G=\mathbf{R}^n$; the general case is argued similarly). The convolution $\tau_1*\tau_2$ of two distributions τ_1, τ_2 is well defined as a distribution provided at most one of them has noncompact support; moreover, the associative law $(\tau_1 * \tau_2) * \tau_3 = \tau_1 * (\tau_2 * \tau_3)$ holds when at most one of the τ_j's has noncompact support. To define convolutions of distributions with noncompact support one must impose additional regularity assumptions, and the associative law need not hold even when all convolutions in question are well defined. We shall therefore need to establish the absence of pathology in certain situations concerning homogeneous kernels.

(1.12) **Lemma.** *Suppose* $0 \leqq \alpha < Q$, $p \geqq 1$, $q > 1$, *and* $r^{-1} = p^{-1} + q^{-1} - (\alpha/Q) - 1 > 0$. *If K is a kernel of type α, $f \in L^p$, and $g \in L^q$,then $f*(g*K)$ and $(f*g)*K$ are well defined as elements of L^r, and they are equal.*

Proof. By Propositions (1.9) and (1.11) and Young's inequality, the mappings $(f, g) \to f*(g*K)$ and $(f, g) \to (f*g)*K$ are continuous from $L^p \times L^q$ to L^r. They coincide when f and g have compact support, and hence in general.

(1.13) **Proposition.** *Suppose K_α is a kernel of type α and K_β is a kernel of type β where $\alpha > 0$, $\beta \geqq 0$, and $\alpha + \beta < Q$. Then $K_\alpha * K_\beta$ is well defined as a kernel of type $\alpha + \beta$. Moreover, if $f \in L^p$ where $1 < p < Q/(\alpha+\beta)$, then $(f*K_\alpha)*K_\beta$ and $f*(K_\alpha * K_\beta)$ belong to L^q, $q^{-1} = p^{-1} - (\alpha+\beta)/Q$, and they are equal.*

Proof. By Proposition (1.4), given $x \neq 0$ we may choose $\varepsilon > 0$ so small that $\{y : |y| < \varepsilon \text{ and } |xy^{-1}| < \varepsilon\}$ is empty. Then if $\beta > 0$,

$$K_\alpha * K_\beta(x) = \Big[\int_{|y|<\varepsilon} + \int_{|xy^{-1}|<\varepsilon} + \int_{|y|>\varepsilon,\, |xy^{-1}|>\varepsilon}\Big] K_\alpha(xy^{-1}) K_\beta(y)\, dy.$$

By Corollary (1.7), these integrals are absolutely convergent since the integrand is $O(|y|^{\beta-Q})$ near 0, $O(|xy^{-1}|^{\alpha-Q})$ near x, and $O(|y|^{\alpha+\beta-2Q})$ near infinity. Likewise, if $\beta = 0$, by Proposition (1.8) we may assume K_β contains no delta function and we have

$$K_\alpha * K_\beta(x) = \int_{|y|<\varepsilon} [K_\alpha(xy^{-1}) - K_\alpha(x)] K_\beta(y)\, dy$$
$$+ \Big[\int_{|xy^{-1}|<\varepsilon} + \int_{|y|>\varepsilon,\, |xy^{-1}|>\varepsilon}\Big] K_\alpha(xy^{-1}) K_\beta(y)\, dy$$

where the first integrand is $O(|y|^{1-Q})$, and again the integrals are absolutely convergent. Thus $K_\alpha * K_\beta(x)$ is well defined for $x \neq 0$, and a simple change of variables shows that $K_\alpha * K_\beta$ is homogeneous of degree $\alpha + \beta - Q$. Moreover, let us choose $\varphi_1 \in \mathscr{D}$ with $\varphi_1(y) = 1$ for $|y| < \varepsilon/2$ and $\varphi_1(y) = 0$ for $|y| > \varepsilon$, and set $\varphi_2(y) = \varphi_1(xy^{-1})$. Then φ_1 and φ_2 have disjoint support, and for $\beta > 0$ we can write $K_\alpha * K_\beta(z) = I_1 + I_2 + I_3$, where

$$I_1 = \int \varphi_1(y) K_\alpha(zy^{-1}) K_\beta(y)\, dy,$$
$$I_2 = \int \varphi_2(y) K_\alpha(zy^{-1}) K_\beta(y)\, dy = \int \varphi_2(y^{-1}z) K_\alpha(y) K_\beta(y^{-1}z)\, dy,$$
$$I_3 = \int [1 - \varphi_1(y) - \varphi_2(y)] K_\alpha(zy^{-1}) K_\beta(y)\, dy,$$

with a similar formula for $\beta=0$. If $|xz^{-1}|<\varepsilon/2$, the factor of the integrand containing z in each of these integrals is $\mathscr{C}^\infty$. Also, every derivative is the sum of homogeneous terms of positive degree, so the derived integrand in I_3 remains $O(|y|^{\alpha+\beta-2Q})$. It follows that $K_\alpha * K_\beta$ is $\mathscr{C}^\infty$ away from 0, and hence is a kernel of type $\alpha+\beta$.

Next, if $f\in L^p$ where $p>1$ and $q^{-1}=p^{-1}-(\alpha+\beta)/Q>0$, we observe that $(f*K_\alpha)*K_\beta$ and $f*(K_\alpha*K_\beta)$ are in L^q by Propositions (1.9) and (1.11) and Young's inequality. To complete the proof, it suffices to show that $(f*K_\alpha)*K_\beta$ and $f*(K_\alpha*K_\beta)$ are equal as distributions. Define $K_\alpha^0(x)=K_\alpha(x)$ if $|x|\leqq 1$ and $=0$ otherwise, and set $K_\alpha^\infty=K_\alpha-K_\alpha^0$. By Corollary (1.7), if $r=Q/(Q-\alpha)$, $K_\alpha^0\in L^{r-\varepsilon}$ and $K_\alpha^\infty\in L^{r+\varepsilon}$ for any $\varepsilon>0$. Taking ε so small that $r-\varepsilon>1$ and $p^{-1}+(r+\varepsilon)^{-1}-(\beta/Q)-1>0$, by Lemma (1.12) we see that $(f*K_\alpha^0)*K_\beta$ and $f*(K_\alpha^0*K_\beta)$ coincide as elements of L^s where $s^{-1}=p^{-1}+(r-\varepsilon)^{-1}+(\beta/Q)-1$, and $(f*K_\alpha^\infty)*K_\beta$ and $f*(K_\alpha^\infty*K_\beta)$ coincide as elements of L^t where $t^{-1}=p^{-1}+(r+\varepsilon)^{-1}+(\beta/Q)-1$. Thus $(f*K_\alpha)*K_\beta$ and $f*(K_\alpha*K_\beta)$ coincide as elements of L^s+L^t, and we are done.

We shall occasionally wish to use the additive structure on G defined by $x+y=\exp(\exp^{-1}x+\exp^{-1}y)$. We note that dilations distribute over addition: $\gamma_r(x+y)=\gamma_r x+\gamma_r y$.

(1.14) **Lemma.** *Given a fixed $x\in G$, define $\varphi\colon G\to G$ by $\varphi(y)=xy-x$. Then $\|\varphi(y)\|=O(\|y\|)$ and $\|\varphi(y^{-1})-\varphi(y)^{-1}\|=O(\|y\|^2)$ as $y\to 0$.*

Proof. Let $x=\exp X$, $y=\exp Y$. By the Campbell—Hausdorff formula (cf. Hochschild [9]),

$$\varphi(y)=\exp\left(Y+\frac{1}{2}[X,Y]+\dots\right)$$

where the dots indicate higher order commutators of X and Y, which are finite in number by nilpotency. The assertions are then clear, taking account of the fact that $(\exp Z)^{-1}=\exp(-Z)$ for all $Z\in\mathfrak{g}$.

We now prove a mean-value theorem for homogeneous functions.

(1.15) **Proposition.** *Let f be a homogeneous function of degree λ ($\lambda\in\mathbf{R}$) which is $\mathscr{C}^2$ away from 0. There are constants C, $\varepsilon>0$ such that*

$$|f(xy)-f(x)|\leqq C|y|\,|x|^{\lambda-1}\quad \textit{whenever}\quad |y|\leqq\frac{1}{2}|x|,$$

$$|f(xy)+f(xy^{-1})-2f(x)|\leqq C|y|^2|x|^{\lambda-2}\quad\textit{whenever}\quad |y|\leqq\varepsilon|x|.$$

Proof.. If x and y are replaced by $\gamma_r x$ and $\gamma_r y$, both sides of both inequalities are multiplied by r^λ, so it suffices to assume $|x|=1$ and $|y|\leqq\frac{1}{2}$ or $|y|\leqq\varepsilon$. If $|x|=1$ and $|y|\leqq\frac{1}{2}$ then xy is bounded away from 0, so since the mapping $y\to xy$ is smooth,

$$|f(xy)-f(x)|\leqq C\|y\|=C\|y\|\,|x|^{\lambda-1}$$

with C independent of x, y in the given regions. The first assertion then follows from Lemma (1.3). Moreover, if we set $z=\varphi(y)$ as in Lemma (1.14), we can choose $\varepsilon>0$ small enough so that $|z|\leqq\frac{1}{2}$ whenever $|y|\leqq\varepsilon$ and $|x|=1$. As above we have

$$|f(x+z)+f(x-z)-2f(x)| \leqq C\|z\|^2.$$

But then by Lemma (1.14),

$$\begin{aligned}|f(xy)+f(xy^{-1})-2f(x)| &= |f(\varphi(y))+f(\varphi(y^{-1}))-2f(x)|\\ &\leqq |f(x+z)+f(x-z)-2f(x)|+|f(\varphi(y)^{-1})-f(\varphi(y^{-1}))|\\ &\leqq C\|z\|^2+C\|\varphi(y)^{-1}-\varphi(y^{-1})\|\\ &\leqq C\|z\|^2+C\|y\|^2\leqq C'\|y\|^2\leqq C''|y|^2\\ &\leqq C''|y|^2|x|^{\lambda-2}.\end{aligned}$$

We shall occasionally need to consider right-invariant derivatives as well as left-invariant ones. If X is a left-invariant vector field, we shall denote by $\tilde{X}$ the right-invariant vector field which agrees with X at 0: if $J(x)=x^{-1}$, we have $\tilde{X}=-J_*X$. Since J commutes with dilations, it is clear that $\tilde{X}$ is homogeneous of degree λ if and only if X is.

In the later parts of this paper we will be concerned with an important class of groups with dilations, the "stratified" groups. If $\mathfrak{g}$ is a nilpotent Lie algebra, a *stratification* of $\mathfrak{g}$ is a decomposition of $\mathfrak{g}$ as a vector space sum, $\mathfrak{g}=V_1\oplus\ldots\oplus V_m$, such that $[V_1, V_j]=V_{j+1}$ for $1\leqq j<m$ and $[V_1, V_m]=\{0\}$. We note that the stratification is completely determined by V_1, and that $\oplus_k^m V_j$ is the ideal of the k-th order commutators. If $\mathfrak{g}$ is stratified, it admits a canonical family of dilations, namely

$$\gamma_r(X_1+X_2+\ldots+X_m) = rX_1+r^2X_2+\ldots+r^mX_m \quad (X_j\in V_j). \tag{1.16}$$

We note that the homogeneous dimension of $\mathfrak{g}$ is $\sum_1^m j(\dim V_j)$. Having chosen a Euclidean norm $\|\cdot\|$ on $\mathfrak{g}$ with respect to which the V_j's are mutually orthogonal, we define a homogeneous norm on the corresponding group G by

$$\left|\exp\left(\sum_1^m X_j\right)\right| = \left(\sum_1^m \|X_j\|^{2m!/j}\right)^{1/2m!} \quad (X_j\in V_j). \tag{1.17}$$

Henceforth, by a *stratified group* we shall mean a simply connected nilpotent group G together with a stratification $\mathfrak{g}=\oplus_1^m V_j$ of its Lie algebra and the dilations and homogeneous norm defined by (1.16) and (1.17).

Here are some examples of Lie algebras with stratifications:

(a) $\mathfrak{g}$ Abelian, $m=1$, $V_1=\mathfrak{g}$.

(b) $\mathfrak{g}$ any algebra of step two (i.e., $[\mathfrak{g},\mathfrak{g}]\neq\{0\}$ but $[[\mathfrak{g},\mathfrak{g}],\mathfrak{g}]=\{0\}$), $m=2$, $V_1=$any subspace complementary to $[\mathfrak{g},\mathfrak{g}]$.

(c) $\mathfrak{g}=$the quotient of the free Lie algebra $\mathfrak{g}_n$ on n generators $X_1, \ldots, X_n$ by the ideal $[[\ldots[\mathfrak{g}_n, \mathfrak{g}_n], \ldots \mathfrak{g}_n], \mathfrak{g}_n]$ ($m+1$ factors), $V_1=$the linear span of $X_1, \ldots, X_n$.

(d) $\mathfrak{g}$=the algebra of $(m+1)\times(m+1)$ real or complex matrices (a_{ij}) such that $a_{ij}=0$ for $i\geqq j$ (with Lie product $[A, B]=AB-BA$), $V_1=\{(a_{ij}): a_{ij}=0$ for $i\neq j-1\}$.

(e) $\mathfrak{g}$=the nilpotent part of the Iwasawa decomposition of a semisimple Lie algebra, V_1=the linear span of the root spaces of the simple roots.

Let G be stratified. It is clear that $X\in\mathfrak{g}$ is homogeneous of degree j if and only if $X\in V_j$. We choose once and for all a basis $X_1, \dots, X_n$ for V_1 and set $\mathscr{J}=-\sum_1^n X_j^2$. $\mathscr{J}$ is a left-invariant second-order differential operator which is homogeneous of degree 2; we call it the *sub-Laplacian* of G (relative to the stratification and the basis $X_1, \dots, X_n$). As we shall see, $\mathscr{J}$ plays much the same role on G as (minus) the ordinary Laplacian does on $\mathbf{R}^n$.

Henceforth, if G is a nilpotent group with dilations $\{\gamma_r\}$, we shall generally denote $\gamma_r x$ simply by rx.

2. Fundamental solutions for homogeneous operators

In this section we assume G is a nilpotent group with dilations, of homogeneous dimension Q. We recall that if D is a differential operator, its transpose D^t is defined by $\int(D^t u)v=\int u(Dv)$ for all $u, v\in\mathscr{D}$. Our main result is the following:

(2.1) **Theorem.** *Let $\mathscr{L}$ be a homogeneous differential operator on G of degree α, $0<\alpha<Q$, such that $\mathscr{L}$ and $\mathscr{L}^t$ are both hypoelliptic. Then there is a unique kernel K_0 of type α which is a fundamental solution for $\mathscr{L}$ at 0, i.e., which satisfies $\mathscr{L}K_0=\delta$.*

The main tool in the proof is the following theorem from functional analysis, which combines Theorems 52.1 and 52.2 of Trèves [29]:

(2.2) **Lemma.** *Let D be a differential operator on a Euclidean space $\mathbf{R}^N$ such that D and D^t are both hypoelliptic. Then for each $x\in\mathbf{R}^N$ there is an open neighborhood U of x and a distribution $K\in\mathscr{D}'(U)$ which is $\mathscr{C}^\infty$ away from x such that $DK(y)=\delta(y-x)$ on U. Moreover, the topologies of $\mathscr{E}(U)$ and $\mathscr{D}'(U)$ coincide on $\mathscr{N}=\{f\in\mathscr{D}'(U): Df=0\}$ and make $\mathscr{N}$ into a Frèchet space.*

Proof of Theorem (2.1). We apply Lemma (2.2) with $x=0$ and $D=\mathscr{L}$ to obtain a neighborhood U of 0 and a distribution $K\in\mathscr{D}'(U)$ which is $\mathscr{C}^\infty$ away from 0 such that $\mathscr{L}K=\delta$ on U. By shrinking U if necessary, we may assume that $U=\{x: |x|<C\}$, so that $x\in U$ implies $rx(=\gamma_r x)\in U$ for $r\leqq 1$. Then for $0<r\leqq 1$ we define the distribution $h_r\in\mathscr{D}'(U)$ by

$$h_r = K-r^{Q-\alpha}(K\circ\gamma_r).$$

By the homogeneity of $\mathscr{L}$ we have

$$\mathscr{L}h_r = \delta-r^Q(\delta\circ\gamma_r) = \delta-\delta = 0,$$

so that $h_r \in \mathcal{N} \subset \mathcal{E}(U)$. Next, observe that if $s \leqq r$,

$$h_s(x) - h_r(x) = r^{Q-\alpha}K(rx) - s^{Q-\alpha}K(sx) \tag{2.3}$$
$$= r^{Q-\alpha}[K(rx) - (s/r)^{Q-\alpha}K((s/r)rx)]$$
$$= r^{Q-\alpha}h_{s/r}(rx).$$

If we set $s = r^2$ in (2.3), we obtain

$$h_{r^2}(x) = r^{Q-\alpha}h_r(rx) + h_r(x). \tag{2.4}$$

Replacing r by r^2 in (2.4) and substituting (2.4) in the result,

$$h_{r^4}(x) = r^{2(Q-\alpha)}h_{r^2}(r^2x) + h_{r^2}(x)$$
$$= r^{3(Q-\alpha)}h_r(r^3x) + r^{2(Q-\alpha)}h_r(r^2x) + r^{Q-\alpha}h_r(rx) + h_r(x).$$

Continuing inductively, we obtain

$$h_{r^{2^n}}(x) = \sum_{k=0}^{2n-1} r^{k(Q-\alpha)}h_r(r^kx). \tag{2.5}$$

If we set $V_\varepsilon = \{x : |x| \leqq C - \varepsilon\} \subset U$, (2.5) yields

$$\sup_{x \in V_\varepsilon} |h_{r^{2^n}}(x)| \leqq (1 - r^{Q-\alpha})^{-1} \sup_{x \in V_\varepsilon} |h_r(x)| \tag{2.6}$$

for all n. Now $r \to h_r$ is clearly continuous from (0, 1] to $\mathcal{D}'(U)$, so $\{h_r : \frac{1}{4} \leqq r \leqq \frac{1}{2}\}$ is compact in $\mathcal{D}'(U)$, hence in $\mathcal{E}(U)$. But any $s \leqq \frac{1}{2}$ can be expressed as $s = r^{2^n}$ for some n and some $r \in [\frac{1}{4}, \frac{1}{2}]$ so that by (2.6),

$$\sup_{\substack{x \in V_\varepsilon \\ s \leqq 1}} |h_s(x)| \leqq (1 - 2^{\alpha - Q})^{-1} \sup_{\substack{x \in V_\varepsilon \\ \frac{1}{4} \leqq r \leqq \frac{1}{2}}} |h_r(x)| + \sup_{\substack{x \in V_\varepsilon \\ \frac{1}{2} \leqq s \leqq 1}} |h_s(x)| = C_\varepsilon < \infty.$$

Thus the h_r's are uniformly bounded on V_ε. But if $s < r$ and $x \in V_\varepsilon$, (2.3) implies

$$|h_s(x) - h_r(x)| \leqq r^{Q-\alpha}|h_{s/r}(rx)| \leqq C_\varepsilon r^{Q-\alpha} \to 0 \quad \text{as} \quad r, s \to 0.$$

Thus the h_r's are uniformly Cauchy on compact subsets of U as $r \to 0$, so they are Cauchy in $\mathcal{D}'(U)$, hence in $\mathcal{E}(U)$, and the limit h_0 satisfies $\mathcal{L}h_0 = 0$. Now set

$$K_0 = K - h_0 = \lim_{r \to 0} r^{Q-\alpha}(K \circ \gamma_r).$$

On the one hand, $\mathcal{L}K_0 = \mathcal{L}K - \mathcal{L}h_0 = \delta$, and on the other, if $0 < s < 1$,

$$K_0(sx) = \lim_{r \to 0} r^{Q-\alpha}K(srx) = \lim_{r \to 0} (r/s)^{Q-\alpha}K(rx) = s^{\alpha - Q}K_0(x). \tag{2.7}$$

But now we can extend K_0 to the whole space by requiring (2.7) to hold for all $s > 0$, so that K_0 is a kernel of type α, and the homogeneity of $\mathcal{L}$ guarantees that the equation $\mathcal{L}K_0 = \delta$ holds globally.

Finally, if K_1 were another kernel of type α satisfying $\mathscr{L}K_1=\delta$, then K_0-K_1 would be $\mathscr{C}^\infty$ even at 0 since $\mathscr{L}(K_0-K_1)=0$. Since K_0-K_1 is homogeneous of degree $\alpha-Q<0$, we must have $K_0-K_1=0$. The proof is complete.

Remark. This theorem applies equally well to $\mathscr{L}^t$, and the corresponding kernel K_0^t is given by $K_0^t(x)=K_0(x^{-1})$.

(2.8) **Corollary.** *Let $\mathscr{L}$, $\mathscr{L}^t$ be as in Theorem* (2.1), *and K_0, K_0^t their fundamental solutions. If in addition $\mathscr{L}$ is left-invariant, then for any $\tau\in\mathscr{E}'$,*

$$\mathscr{L}(\tau*K_0) = \mathscr{L}^t(\tau*K_0^t) = (\mathscr{L}^t\tau)*K_0^t = (\mathscr{L}\tau)*K_0 = \tau.$$

Proof. Since $\mathscr{L}$ and $\mathscr{L}^t$ are left-invariant, $\mathscr{L}(\tau*K_0)=\tau*\mathscr{L}K_0=\tau*\delta=\tau$ and likewise $\mathscr{L}^t(\tau*K_0^t)=\tau$. On the other hand, the mappings $u\to\mathscr{L}(u*K_0)$ and $u\to\mathscr{L}^t(u*K_0^t)$ are *a priori* continuous from $\mathscr{D}$ to $\mathscr{E}$, and their dual mappings from $\mathscr{E}'$ to $\mathscr{D}'$ are $\tau\to(\mathscr{L}^t\tau)*K_0^t$ and $\tau\to(\mathscr{L}\tau)*K_0$, respectively.

(2.9) **Corollary.** *Under the hypotheses of Corollary* (2.8), *for any $\tau\in\mathscr{E}'$ there exists $\sigma\in\mathscr{D}'$ satisfying $\mathscr{L}\sigma=\tau$, and there are no nontrivial solutions in $\mathscr{E}'$ of $\mathscr{L}\tau=0$.*

We now give some examples of Theorem (2.1). The first three are applications of Hörmander's theorem (0.1); we also use the fact that if $X\in\mathfrak{g}$ then $X^t=-X$, since the translations generated by X are isometries.

(2.10) Let G be stratified. Then the sub-Laplacian $\mathscr{J}$ is homogeneous of degree 2 and hypoelliptic, and $\mathscr{J}^t=\mathscr{J}$, so Theorem (2.1) applies provided $Q>2$. In particular, if $G=\mathbf{R}^n$, $\mathscr{J}=-\sum_1^n(\partial/\partial x_j)^2$, and $|x|=(\sum_1^n x_j^2)^{1/2}$, then the fundamental solution for $\mathscr{J}$ is of course

$$K_0(x) = \frac{\Gamma(n/2)}{2\pi^{n/2}(n-2)}|x|^{2-n} \quad (n\neq 2).$$

However, in case $n=2$, the fundamental solution for $\mathscr{J}$ is $(2\pi)^{-1}\log(|x|^{-1})$, which illustrates how Theorem (2.1) can break down when $\alpha\geqq Q$.

(2.11) Let G be stratified, and let $\tilde{G}=G\times\mathbf{R}$. We define dilations on $\tilde{G}$ by $\gamma_r(x,t)=$ $=(rx, r^2t)$. Then the "heat operator" $\mathscr{J}+(\partial/\partial t)$ and its transpose $\mathscr{J}-(\partial/\partial t)$ are hypoelliptic and homogeneous of degree $2(<Q)$. Again, in the case $G=\mathbf{R}^n$, the fundamental solution is well known to be

$$K_0(x,t) = \begin{cases}(4\pi t)^{-n/2}e^{-|x|^2/4t} & (t>0)\\ 0 & (t\leqq 0).\end{cases}$$

(2.12) Let G be stratified and non-Abelian. If Y is any element of $V_2\subset\mathfrak{g}$, then $\mathscr{J}+Y$ and its transpose $\mathscr{J}-Y$ are hypoelliptic and homogeneous of degree $2(<Q)$.

(2.13) Let G be stratified and non-Abelian. Suppose $T\in V_2$ is such that there exists a basis $X_1, \dots, X_n$, $Y_1, \dots, Y_n$, $W_1, \dots, W_k$ of V_1 with $T=\sum_1^n [Y_j, X_j]$. For $\alpha\in\mathbf{C}$, set

$$\mathcal{L}_\alpha = -\sum_1^n (X_j^2+Y_j^2) - \sum_1^k W_j^2 + \alpha T.$$

Then $\mathcal{L}_\alpha$ and its transpose $\mathcal{L}_{-\alpha}$ are homogeneous of degree $2(<Q)$, and we claim $\mathcal{L}_\alpha$ is hypoelliptic provided $|\mathrm{Im}\,\alpha|<1$. Hörmander's theorem does not apply unless $\mathrm{Im}\,\alpha=0$, so we use the method of L^2 estimates.

First we note that for any $u\in\mathcal{D}$,

$$(\mathcal{L}_0 u, u) = \sum_1^n (\|X_j u\|_2^2 + \|Y_j u\|_2^2) + \sum_1^k \|W_j u\|_2^2,$$

so by a well-known estimate (cf. Theorem 5.4.7 of Folland—Kohn [5], or Oleĭnik—Radkevič [22]), of which we shall prove a sharper version later (Theorem (4.16)), there exist $C, \varepsilon>0$ such that

$$\|u\|_{(2,\varepsilon)}^2 \leqq C[(\mathcal{L}_0 u, u) + \|u\|_2^2] \quad (u\in\mathcal{D}), \tag{2.14}$$

where $\|\ \|_{(2,\varepsilon)}$ is the L^2 Sobolev norm of order ε. We next show that

$$|(iTu, u)| \leqq (\mathcal{L}_0 u, u). \tag{2.15}$$

Indeed, set $Z_j=2^{-1/2}(X_j-iY_j)$ and $\bar{Z}_j=2^{-1/2}(X_j+iY_j)$. Then $[\bar{Z}_j, Z_j]=i[Y_j, X_j]$, so that

$$|(iTu, u)| = \left|\sum_1^n ([\bar{Z}_j, Z_j]u, u)\right| \leqq \sum_1^n (|(\bar{Z}_j Z_j u, u)| + |(Z_j\bar{Z}_j u, u)|)$$
$$= \sum_1^n (\|Z_j u\|_2^2 + \|\bar{Z}_j u\|_2^2) = \sum_1^n (\|X_j u\|_2^2 + \|Y_j u\|_2^2) \leqq (\mathcal{L}_0 u, u).$$

Thirdly, we note that

$$\mathrm{Re}\,(\mathcal{L}_\alpha u, u) = (\mathcal{L}_0 u, u) + (\mathrm{Re}\,\alpha)\,\mathrm{Re}\,(Tu, u) - (\mathrm{Im}\,\alpha)\,\mathrm{Re}\,(iTu, u).$$

But $\mathrm{Re}\,(Tu, u)=0$ since T is skew-symmetric, so by (2.15),

$$\mathrm{Re}\,(\mathcal{L}_\alpha u, u) \geqq (1-|\mathrm{Im}\,\alpha|)(\mathcal{L}_0 u, u). \tag{2.16}$$

(2.14, 15, 16) combined then yield (for $|\mathrm{Im}\,\alpha|<1$)

$$\|u\|_{(2,\varepsilon)}^2 \leqq C(1-|\mathrm{Im}\,\alpha|)^{-1}(\mathrm{Re}\,(\mathcal{L}_\alpha u, u) + \|u\|_2^2) \tag{2.17}$$

and

$$|\mathrm{Im}\,(\mathcal{L}_\alpha u, u)| = |\mathrm{Re}\,\alpha|\,|(iTu, u)| \leqq (1-|\mathrm{Im}\,\alpha|)^{-1}|\mathrm{Re}\,\alpha|\,\mathrm{Re}\,(\mathcal{L}_\alpha u, u). \tag{2.18}$$

But by the Kohn—Nirenberg regularity theorem [14], (2.17) and (2.18) imply that $\mathcal{L}_\alpha$ is hypoelliptic for $|\mathrm{Im}\,\alpha|<1$.

A particular case of this construction is of interest in the theory of several complex variables. Let $G=H_n\times\mathbf{R}^k$, where H_n is the Heisenberg group of dimension

$2n+1$. Let $\tilde{X}_1, \ldots, \tilde{X}_n, \tilde{Y}_1, \ldots, \tilde{Y}_n, \tilde{T}$ be a basis for the Lie algebra of H_n satisfying the canonical commutation relations

$$[\tilde{Y}_j, \tilde{X}_j] = \tilde{T} \quad \text{for} \quad j = 1, \ldots, n; \quad \text{all other brackets} = 0,$$

and let $\tilde{W}_1, \ldots, \tilde{W}_n$ be a basis for the Lie algebra of $\mathbf{R}^k$. Given positive numbers $a_1, \ldots, a_n, b_1, \ldots, b_n$ and a complex number β, set

$$\mathscr{L}_\beta = -\sum_1^n (a_j^2 X_j^2 + b_j^2 Y_j^2) - \sum_1^k \tilde{W}_j^2 + \beta \tilde{T}.$$

Applying the above result with $X_j = a_j \tilde{X}_j$, $Y_j = b_j \tilde{Y}_j$, $W_j = \tilde{W}_j$, and $T = (\sum_1^n a_j b_j)\tilde{T}$, we see that $\mathscr{L}_\beta$ is hypoelliptic provided $|\operatorname{Im} \beta| < \sum_1^n a_j b_j$.

In case $k = 2m$ is even, G can be imbedded in a natural way in $\mathbf{C}^{n+m+1}$ as a real hypersurface whose Levi form has n non-zero eigenvalues at each point, and the operators $\mathscr{L}_\beta$, for various imaginary values of β, are closely related to the "Laplacian" $\Box_b$ of the tangential $\bar{\partial}$ complex on $G \subset \mathbf{C}^{n+m+1}$; cf. Folland—Stein [7]. In particular if $a_j = b_j = 1$ for all j and $k = 0$, it is shown in [7] that $\mathscr{L}_\beta$ is hypoelliptic unless $\pm i\beta = n, n+2, n+4, \ldots$, and the fundamental solution for $\mathscr{L}_\beta$ is computed explicitly.

We conclude this section with a technical result that will be useful later.

(2.19) **Proposition.** *Let G be stratified and of homogeneous dimension $Q > 2$, and let $\mathscr{J} = -\sum_1^n X_j^2$ be a sub-Laplacian on G. Then $\mathscr{J}(\mathscr{D})$ is dense in L^p, $1 < p < \infty$.*

Proof. Since $p < \infty$, it suffices to show that $\mathscr{D}$ is in the closure of $\mathscr{J}(\mathscr{D})$ in the L^p norm. Let K_0 be the fundamental solution for $\mathscr{J}$ given by Theorem (2.1), and choose $\varphi \in \mathscr{D}$ such that $\varphi(x) = 1$ when $|x| \leqq 1$ and $\varphi(x) = 0$ when $|x| \geqq 2$. Given $f \in \mathscr{D}$, set $u = f * K_0$ and $u_k(x) = \varphi(2^{-k}x)u(x)$ $(k = 1, 2, 3, \ldots)$. Then $u_k \in \mathscr{D}$, and we claim that $\mathscr{J}u_k \to f$ in the L^p norm provided $p > 1$. Indeed, by Corollary (2.8) and the homogeneity of $\mathscr{J}$,

$$\mathscr{J}u_k(x) = \varphi(2^{-k}x)f(x) + 2^{-2k}(\mathscr{J}\varphi)(2^{-k}x)u(x) - 2\sum_1^n 2^{-k}[(X_j\varphi)(2^{-k}x)][(X_j u)(x)].$$

Since $\varphi(2^{-k}x)f(x) = f(x)$ for sufficiently large k, we must show that the other terms tend to zero.

Since $K_0(x) = O(|x|^{2-Q})$ as $x \to \infty$, the same is true of $u(x) = (f * K_0)(x)$. Likewise, $X_j K_0(x) = O(|x|^{1-Q})$ and so $X_j u(x) = (f * X_j K_0)(x) = O(|x|^{1-Q})$. Thus by Corollary (1.6),

$$\int |2^{-2k}(\mathscr{J}\varphi)(2^{-k}x)u(x)|^p \, dx \leqq C 2^{-2kp} \|\mathscr{J}\varphi\|_\infty^p \int_{2^k \leqq |x| \leqq 2^{k+1}} |x|^{p(2-Q)} \, dx$$

$$\leqq C' \|\mathscr{J}\varphi\|_\infty^p 2^{-kQ(p-1)}$$

which tends to zero as $k\to\infty$ since $p>1$. Likewise,

$$\int |2^{-k}[(X_j\varphi)(2^{-k}x)][(X_ju)(x)]|^p\,dx \leqq C\|X_j\varphi\|_\infty^p 2^{-kp}\int_{2^k\leqq|x|\leqq 2^{k+1}}|x|^{p(1-Q)}\,dx$$
$$\leqq C'\|X_j\varphi\|_\infty^p 2^{-kQ(p-1)},$$

and the proof is complete.

Remark. The same reasoning shows that $\mathcal{J}(\mathcal{D})$ is dense in $\mathcal{C}_0$ in the uniform norm. $\mathcal{J}(\mathcal{D})$ is not dense in L^1, however, for every $f\in\mathcal{J}(\mathcal{D})$ satisfies $\int f(x)\,dx=0$.

3. Analysis of the sub-Laplacian

Henceforth we assume that G is a stratified group of homogeneous dimension $Q>2$. (The latter requirement excludes only $\mathbf{R}^1$ and $\mathbf{R}^2$, for which our major results are already well known.) The purpose of this section is to develop the theory of complex powers of the sub-Laplacian $\mathcal{J}=-\sum_n^1 X_j^2$ on G. The principal tool for this purpose is the diffusion semigroup H_t generated by $-\mathcal{J}$, whose principal properties are summarized in the following theorem.

(3.1) **Theorem.** *There is a unique semigroup $\{H_t\colon 0<t<\infty\}$ of linear operators on L^1+L^∞ satisfying the following conditions:*

(i) *$H_t f=f*h_t$ where $h_t(x)=h(x,t)$ is $\mathcal{C}^\infty$ on $G\times(0,\infty)$, $\int h_t(x)\,dx=1$ for all t, and for all x and t, $h(x,t)\geqq 0$ and*

(3.2) $$h(rx, r^2t) = r^{-Q}h(x,t).$$

(ii) *If $u\in\mathcal{D}$, $\lim_{t\to 0}\|t^{-1}(H_tu-u)+\mathcal{J}u\|_\infty = 0$.*

Moreover, $\{H_t\}$ has the following properties:

(iii) *$\{H_t\}$ is a contraction semigroup on L^p, $1\leqq p\leqq\infty$, which is strongly continuous for $p<\infty$. Also, if $1<p<\infty$, $\{H_t\}$ can be extended to a holomorphic contraction semigroup $\{H_z\colon|\arg z|<\frac{1}{2}\pi(1-|1-(2/p)|)\}$ on L^p.*

(iv) *H_t is self-adjoint, i.e., $H_t|L^p$ is the dual of $H_t|L^{p'}$ where $p^{-1}+(p')^{-1}=1$, $p>1$.*

(v) *$f\geqq 0$ implies $H_tf\geqq 0$, and $H_t1=1$.*

Proof. Let $\tilde{\mathcal{D}}$ be the space of $\mathcal{C}^\infty$ functions which are constant outside a compact set, and let $Y_1,\dots,Y_N$ be a basis for the Lie algebra of G. Let $\tilde{\mathcal{C}}^2$ (resp. $\tilde{\mathcal{C}}$) be the completion of $\tilde{\mathcal{D}}$ with respect to the norm

$$|||f||| = \|f\|_\infty+\sum_{j=1}^N\|Y_jf\|_\infty+\sum_{j,k=1}^N\|Y_jY_kf\|_\infty$$

(resp. the uniform norm). According to a theorem of G. Hunt [11] there is a unique strongly continuous semigroup $\{H_t\}$ on $\tilde{\mathcal{C}}$ such that

(a) for each $t>0$ there is a probability measure μ_t on G such that $H_t f(x)= =\int f(xy^{-1})\,d\mu_t(y)$;

(b) the infinitesimal generator of $\{H_t\}$ is defined on $\tilde{\mathscr{C}}^2$ and coincides with $-\mathscr{J}$ there.

Moreover, $\lim_{t\to 0}\mu_t(E)=1$ whenever $0\in E\subset G$, and (since $\mathscr{J}$ is symmetric) $d\mu_t(y)=d\mu_t(y^{-1})$. We also note that since $\mathscr{J}$ annihilates constants and $\tilde{\mathscr{D}}$ is dense in $\tilde{\mathscr{C}}^2$, the action of $\mathscr{J}$ on $\mathscr{D}$ determines $\{H_t\}$.

Let h be the distribution on $G\times(0,\infty)$ defined by

$$\langle h, u\otimes v\rangle = \int_0^\infty\int_G u(x)v(t)\,d\mu_t(x)\,dt \quad (u\in\mathscr{D}(G),\ v\in\mathscr{D}((0,\infty))).$$

Then because of (b) it is easily verified that

$$\langle h, (\mathscr{J}u)\otimes v\rangle = \langle h, u\otimes(dv/dt)\rangle$$

so that h is a distribution solution $(\mathscr{J}+(\partial/\partial t))h=0$. But by Hörmander's theorem (0.1), $\mathscr{J}+(\partial/\partial t)$ is hypoelliptic, so $h\in\mathscr{C}^\infty(G\times(0,\infty))$, and we have $d\mu(x)=h(x,t)\,dx$. Thus $h(x,t)\geqq 0$, $\int h(x,t)\,dx=1$, and H_t is self-adjoint since $h(x,t)=h(x^{-1},t)$.

Also, since $\mathscr{J}$ is homogeneous of degree 2 we have $(\mathscr{J}(u\circ\gamma_r))\circ\gamma_{1/r}=r^2\mathscr{J}u$. Therefore the semigroup $\{H_{r^2t}\}$ generated by $-r^2\mathscr{J}$ is given by $H_{r^2t}(u)=(H_t(u\circ\gamma_r))\circ\gamma_{1/r}$; that is,

$$\int u(xy^{-1})h(y,r^2t)\,dy = \int u(x(ry^{-1}))h(y,t)\,dy = \int u(xy^{-1})h(r^{-1}y,t)r^{-Q}\,dy.$$

Hence $h(y,r^2t)=r^{-Q}h(r^{-1}y,t)$, so (3.2) holds.

(i), (ii), and (iv) are therefore established, and (v) follows from (i). By (i) and Young's inequality, then, $\{H_t\}$ is a contraction semigroup on L^p, $1\leqq p\leqq\infty$, which is strongly continuous for $p<\infty$ since $h_t\to\delta$ as $t\to 0$. Finally, since H_t is self-adjoint on L^2, we can write $H_t=\int_0^\infty e^{-\lambda t}\,dE(\lambda)$ by the spectral theorem. We then define $H_z=\int_0^\infty e^{-\lambda z}\,dE(\lambda)$ for $|\arg z|<\pi/2$, which proves the second half of (iii) for $p=2$. The cases $p=1$ and $p=\infty$ are trivial, and the general case now follows from the Riesz—Thorin—Stein interpolation theorem. (For the details of this argument, see Stein [24].)

(3.3) **Proposition.** *Extend h to $G\times\mathbf{R}$ by setting $h(x,t)=0$ for $t\leqq 0$. Then $h\in\mathscr{D}'(G\times\mathbf{R})$ and h is a fundamental solution for $\mathscr{J}+(\partial/\partial t)$*

Proof. Since $\int h(x,t)\,dx=1$ for $t>0$, it follows from Fubini's theorem that h is integrable over any region which is bounded in t, so h defines a distribution. Given any $u\in\mathscr{D}(G\times\mathbf{R})$, then, we must show that $(\mathscr{J}+(\partial/\partial t))(u*h)=u$.

For $\varepsilon>0$, set $h^\varepsilon(x,t)=h(x,t)$ if $t>\varepsilon$ and $=0$ otherwise. Then

$$u*h^\varepsilon(x,t) = \int_{-\infty}^{t-\varepsilon}\int_G u(y,s)h(y^{-1}x,t-s)\,dy\,ds.$$

Therefore, if $0<\delta<\varepsilon$,

$$\|u*h^{\varepsilon}-u*h^{\delta}\|_{\infty} \leqq \|u\|_{\infty}\int_{\delta}^{\varepsilon}\int_G h(y,s)\,dy\,ds = (\varepsilon-\delta)\|u\|_{\infty},$$

so that $\{u*h^{\varepsilon}\}$ is uniformly Cauchy as $\varepsilon\to 0$, and its limit is clearly $u*h$. On the other hand, since $(\mathcal{J}+(\partial/\partial t))h=0$ for $t>0$,

$$(\mathcal{J}+(\partial/\partial t))(u*h^{\varepsilon})(x,t) = \int_G u(y,t-\varepsilon)h(y^{-1}x,\varepsilon)\,dy$$

$$= \int [u(y,t-\varepsilon)-u(y,t)]h(y^{-1}x,\varepsilon)\,dy + \int_G u(y,t)h(y^{-1}x,\varepsilon)\,dy = I_1^{\varepsilon}+I_2^{\varepsilon}.$$

Now $\|I_1^{\varepsilon}\|_{\infty}\leqq \sup_{y,t}|u(y,t-\varepsilon)-u(y,t)|\to 0$ as $\varepsilon\to 0$, and I_2^{ε} converges uniformly in x for each t to $u(x,t)$. Since $u\in\mathcal{D}$, the convergence is also uniform in t, so that $\|I_2^{\varepsilon}-u\|_{\infty}\to 0$. This completes the proof.

Remark. In view of (3.2), h is the fundamental solution for $\mathcal{J}+(\partial/\partial t)$ given by Theorem (2.1); cf. (2.11).

(3.4) **Corollary.** *h is $\mathcal{C}^{\infty}$ on $(G\times\mathbf{R})-\{(0,0)\}$. In particular, for each $x\neq 0$, $h(x,t)$ vanishes to infinite order as t decreases to* 0.

(3.5) **Corollary.** *For each $t_0>0$ and positive integer N there is a constant $C>0$ such that $|h(x,t)|\leqq C|x|^{-N}$ for $|x|\geqq 1$ and $t\leqq t_0$.*

Proof. Set $y=|x|^{-1}x$. Then by (3.2), $h(x,t)=|x|^{-Q}h(y,|x|^{-2}t)$, and $\sup_{|y|=1,t\leqq t_0}|h(y,|x|^{-2}t)|=O(|x|^{-N})$ for all N by Corollary (3.4).

(3.6) **Corollary.** *Let D be any left-invariant differential operator on G and k any non-negative integer. Then Corollary* (3.5) *remains valid if h is replaced by $(\partial/\partial t)^k Dh$.*

Proof. D is a sum of homogeneous terms D_j of degree $\lambda_j\geqq 0$. Since $(D_jh)(rx,r^2t)=r^{-Q-\lambda_j}(D_jh)(x,t)$ and $(\partial/\partial t)^k h(rx,r^2t)=r^{-Q-2k}(\partial/\partial t)^k h(x,t)$ the argument of Corollary (3.5) still applies.

Let $\mathcal{J}_p$ be minus the infinitesimal generator of $\{H_t\}$ on L^p. $\mathcal{J}_p$ is a closed operator on L^p whose domain is dense for $p<\infty$; we now wish to identify $\mathcal{J}_p$ more precisely.

(3.7) **Lemma.** *$\mathcal{D}\subset\mathrm{Dom}(\mathcal{J}_p)$, $1\leqq p\leqq\infty$, and $\mathcal{J}_pu=\mathcal{J}u$ for $u\in\mathcal{D}$.*

Proof. The case $p=\infty$ is just Theorem (3.1 (ii)). Suppose $p<\infty$ and $u\in\mathcal{D}$. Since $H_tu\to u$ and $H_t\mathcal{J}u\to\mathcal{J}u$ in L^p as $t\to 0$, and $\mathcal{J}_p$ is closed, it suffices to show that $s^{-1}(H_{t+s}u-H_tu)$ converges in L^p to $-H_t\mathcal{J}u$ as $s\to 0$. But $s^{-1}(h_{t+s}-h_t)$ converges pointwise to $(\partial/\partial t)h_t$, so by Corollary (3.6) and the Lebesgue convergence theorem, $s^{-1}(h_{t+s}-h_t)$ converges in L^1 to $(\partial/\partial t)h_t$. Hence by Young's inequality, $\lim s^{-1}(H_{t+s}u-H_tu)$ exists in the L^p norm, $1\leqq p\leqq\infty$. But the limit in the L^{∞} norm is $-\mathscr{J}_\infty H_t u$, and since $u \in \mathrm{Dom}(\mathscr{J}_\infty)$, general semigroup theory guarantees that $\mathscr{J}_\infty H_t u = H_t \mathscr{J}_\infty u = H_t \mathscr{J} u$.

(3.8) **Theorem.** *$\mathscr{J}_p$ is the maximal restriction of $\mathscr{J}$ to L^p ($1 \le p \le \infty$); that is, $\mathrm{Dom}(\mathscr{J}_p)$ is the set of all $f \in L^p$ such that the distribution derivative $\mathscr{J} f$ is in L^p, and $\mathscr{J}_p f = \mathscr{J} f$. Also, if $p < \infty$, $\mathscr{J}_p$ is the smallest closed extension of $\mathscr{J} | \mathscr{D}$ on L^p.*

Proof. Let p' be the conjugate exponent to p. By Theorem (3.1. (iv)) and Phillips' theorem (cf. Yosida [30]), $\mathscr{J}_{p'}$ is the dual operator of $\mathscr{J}_p$ for $p < \infty$, and $\mathscr{J}_1$ is a restriction of the dual of $\mathscr{J}_\infty$. Hence if $f \in \mathrm{Dom}(\mathscr{J}_p)$ and $u \in \mathscr{D}$, $\int (\mathscr{J}_p f) u = \int f (\mathscr{J}_{p'} u) = \int f(\mathscr{J} u)$ by Lemma (3.7), so $\mathscr{J}_p f = \mathscr{J} f$ in the distribution sense. Thus $\mathrm{Dom}(\mathscr{J}_p) \subset \{f \in L^p : \mathscr{J} f \in L^p\}$. On the other hand, suppose $f \in L^p$ and $\mathscr{J} f \in L^p$, and $u \in \mathscr{D}$. Since $H_s u$ is smooth and rapidly decreasing at infinity for $s > 0$ (Corollary (3.6)), by approximating $H_s u$ by elements of $\mathscr{D}$ it is easy to see that $\int (\mathscr{J} f)(H_s u) = \int f(\mathscr{J} H_s u)$. An application of Fubini's theorem then shows that

$$\int_G \left(\int_0^t H_s \mathscr{J} f(x)\, ds\right) u(x)\, dx = \int_G \mathscr{J} f(x) \left(\int_0^t H_s u(x)\, ds\right) dx$$
$$= \int_G f(x) \left(\int_0^t \mathscr{J} H_s u(x)\, ds\right) dx = -\int_G f(x) \big(H_t u(x) - u(x)\big)\, dx$$
$$= -\int_G \big(H_t f(x) - f(x)\big) u(x)\, dx.$$

Therefore $H_t f - f = -\int_0^t H_s \mathscr{J} f\, ds$, so $\lim_{t \to 0} t^{-1}(H_t f - f)$ exists in L^p and equals $-\mathscr{J} f$, i.e., $f \in \mathrm{Dom}(\mathscr{J}_p)$. This proves the first assertion. For the second, we note that if $p < \infty$, the maximal restriction of $\mathscr{J}$ to $L^{p'}$ (namely $\mathscr{J}_{p'}$) is clearly the dual of the closure of $\mathscr{J} | \mathscr{D}$ on L^p, and it is also the dual of $\mathscr{J}_p$; hence the latter two operators are equal.

(3.9) **Proposition.** *If $1 < p < \infty$, the range of $\mathscr{J}_p$ is dense in L^p and the nullspace of $\mathscr{J}_p$ is $\{0\}$.*

Proof. The first assertion follows from Proposition (2.19) and Lemma (3.7), and the second is then true by duality.

We now pass to the study of complex powers of $\mathscr{J}$. Our definitions are motivated by the following formula, valid for $s > 0$ and $\mathrm{Re}\,\alpha > 0$:

$$s^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} e^{-st}\, dt. \tag{3.10}$$

Also, if $s > 0$, $\mathrm{Re}\,\alpha > 0$, and k is an integer greater than $\mathrm{Re}\,\alpha$,

$$s^{\alpha} = \frac{1}{\Gamma(k-\alpha)} \int_0^\infty t^{k-\alpha-1} s^k e^{-st}\, dt. \tag{3.11}$$

If we formally set $s=\mathcal{J}$, then $e^{-t\mathcal{J}}$ "=" H_t, and

$$\mathcal{J}^{-\alpha} \text{“}=\text{”} \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} H_t\, dt, \quad \mathcal{J}^{\alpha} \text{“}=\text{”} \frac{1}{\Gamma(k-\alpha)} \int_0^\infty t^{k-\alpha-1} \mathcal{J}^k H_t\, dt.$$

We proceed to make these notions precise by defining $\mathcal{J}_p^\alpha$ and $\mathcal{J}_p^{-\alpha}$ as operators on L^p for each p, $1<p<\infty$. We exclude $p=1$ and $p=\infty$ because of various technical difficulties, and because most of the results we ultimately wish to prove are false in these cases anyhow. We first note that by Corollary (3.6), $\mathcal{J}^k h_t \in L^1$ and $(\mathcal{J}^k h)(x, t) = t^{-k-(Q/2)}(\mathcal{J}^k h)(t^{-1/2}x, 1)$, whence $\|\mathcal{J}^k h_t\|_1 = t^{-k}\|\mathcal{J}^k h_1\|_1$. Thus if $f\in L^p$, $\mathcal{J}^k H_t f = f*(\mathcal{J}^k h_t)\in L^p$ and $\|\mathcal{J}^k H_t f\|_p \leqq C\|f\|_p t^{-k}$.

Definition. Suppose that $1<p<\infty$, $\operatorname{Re}\alpha>0$, and k is the smallest integer greater than $\operatorname{Re}\alpha$ (i.e., $k=[\operatorname{Re}\alpha]+1$). The operator $\mathcal{J}_p^\alpha$ is defined by

$$\mathcal{J}_p^\alpha f = \lim_{\varepsilon\to 0} \frac{1}{\Gamma(k-\alpha)} \int_\varepsilon^\infty t^{k-\alpha-1} \mathcal{J}^k H_t f\, dt$$

on the domain of all $f\in L^p$ such that the indicated limit exists in the L^p norm. (By the preceding remarks, the integral is absolutely convergent at infinity.) $\mathcal{J}_p^{-\alpha}$ is defined by

$$\mathcal{J}_p^{-\alpha} f = \lim_{\eta\to\infty} \frac{1}{\Gamma(\alpha)} \int_0^\eta t^{\alpha-1} H_t f\, dt$$

on the domain of all $f\in L^p$ such that the indicated limit exists in the L^p norm.

We shall also have occasion to consider complex powers of $I+\mathcal{J}$, where I is the identity operator. To do this we simply replace $\mathcal{J}$ by $I+\mathcal{J}$ and H_t by the semigroup generated by $-(I+\mathcal{J})$, namely $e^{-t}H_t$. Thus if $\operatorname{Re}\alpha>0$ and $k=[\operatorname{Re}\alpha]+1$, we define $(I+\mathcal{J}_p)^\alpha$ by

$$(I+\mathcal{J}_p)^\alpha f = \lim_{\varepsilon\to 0} \frac{1}{\Gamma(k-\alpha)} \int_\varepsilon^\infty t^{k-\alpha-1} (I+\mathcal{J})^k (e^{-t} H_t f)\, dt$$

on the domain of all $f\in L^p$ such that the indicated limit exists in the L^p norm. Also, we define $(I+\mathcal{J}_p)^{-\alpha}$ for $\operatorname{Re}\alpha>0$ by

$$(I+\mathcal{J}_p)^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} e^{-t} H_t\, dt.$$

Here the integral is obviously absolutely convergent, and $(I+\mathcal{J}_p)^{-\alpha}$ is a bounded operator on L^p.

It remains to consider the case $\operatorname{Re}\alpha=0$. To handle this, we first give an alternative characterization of $\mathcal{J}_p^\alpha$ and $(I+\mathcal{J}_p)^\alpha$ for $p=2$. Namely, let $\mathcal{J}_2=\int_0^\infty \lambda\, dE(\lambda)$ be the spectral resolution of the self-adjoint operator $\mathcal{J}_2$. Then for $\operatorname{Re}\alpha\neq 0$,

(3.12) $$\mathcal{J}_2^\alpha = \int_0^\infty \lambda^\alpha\, dE(\lambda), \quad (I+\mathcal{J}_2)^\alpha = \int_0^\infty (1+\lambda)^\alpha\, dE(\lambda).$$

(The first integral is well defined at $\lambda=0$ since 0 is not an eigenvalue of $\mathscr{J}_2$, cf. Proposition (3.9).) Indeed, the functional calculus provided by the spectral theorem quickly reduces these equations to the formulas (3.10) and (3.11), and equality of domains is easily checked. But (3.12) makes sense even for $\operatorname{Re}\alpha=0$, and so we use it as a definition of $\mathscr{J}_2^\alpha$ and $(I+\mathscr{J}_2)^\alpha$ for all $\alpha\in\mathbf{C}$. To extend this definition to other values of p, we invoke the following multiplier theorem, which is a consequence of Stein's generalized Littlewood—Paley theory.

(3.13) **Lemma.** *Let $\{T_t\}_{t>0}$ be a semigroup on L^1+L^∞ satisfying conditions* (iii), (iv), *and* (v) *of Theorem* (3.1). *Let $-A$ be the infinitesimal generator of $\{T_t\}$ on L^2, and let $A=\int_0^\infty \lambda\, dE(\lambda)$ be its spectral resolution. Moreover, let $\varphi(s)$ be a bounded function on $(0,\infty)$ and let $m(\lambda)=\lambda\int_0^\infty e^{-\lambda s}\varphi(s)\,ds$. Then for $1<p<\infty$ there is a constant C_p, independent of φ, such that $\|m(A)f\|_p\leqq C_p(\sup_{s>0}|\varphi(s)|)\|f\|_p$ for all $f\in L^2\cap L^p$, where $m(A)=\int_{0+}^\infty m(\lambda)\,dE(\lambda)$.*

This lemma is just Corollary 3 on p. 121 of Stein [24]. The dependence of the bound of $m(A)$ on φ is not stated explicitly there, but it follows from the proof.

The definition of $\mathscr{J}_p^\alpha$ and $(I+\mathscr{J}_p)^\alpha$ for $\operatorname{Re}\alpha=0$ is contained in the following proposition.

(3.14) **Proposition.** *If* $\operatorname{Re}\alpha=0$, *$\mathscr{J}_p^\alpha$ and $(I+\mathscr{J}_2)^\alpha$ extend to bounded operators $\mathscr{J}_p^\alpha$ and $(I+\mathscr{J}_p)^\alpha$ on L^p, $1<p<\infty$. Moreover, there is a constant $C_p>0$ such that for all $f\in L^p$,*

$$\|\mathscr{J}_p^\alpha f\|_p\leqq C_p|\Gamma(1-\alpha)|^{-1}\|f\|_p,\quad \|(I+\mathscr{J}_p)^\alpha f\|_p\leqq C_p|\Gamma(1-\alpha)|^{-1}\|f\|_p.$$

Proof. In the terminology of Lemma (3.13), $\mathscr{J}_2^\alpha=m_1(\mathscr{J}_2)$ and $(I+\mathscr{J}_2)^\alpha=m_2(\mathscr{J}_2)$ where $m_1(\lambda)=\lambda^\alpha$ and $m_2(\lambda)=(1+\lambda)^\alpha$. Thus we have merely to observe that

$$\lambda^\alpha=\frac{\lambda}{\Gamma(1-\alpha)}\int_0^\infty e^{-\lambda s}s^{-\alpha}\,ds,$$

$$(1+\lambda)^\alpha=\frac{\lambda}{\Gamma(1-\alpha)}\int_0^\infty e^{-\lambda s}\left[e^{-s}s^{-\alpha}+\int_0^s e^{-\sigma}\sigma^{-\alpha}\,d\sigma\right]ds,$$

and that the absolute values of the integrals are bounded uniformly in α for $\operatorname{Re}\alpha=0$.

In the next theorem we summarize the fundamental properties of the operators $\mathscr{J}_p^\alpha$ which are derived from the general theory of fractional powers of operators. We shall refer to the comprehensive treatment of this subject in the papers of H. Komatsu ([15], [16], [17], [18], [19], [20]); the reader may consult these papers for references to the related works of other authors.

(3.15) **Theorem.** *Let $\mathscr{M}_p$ denote either $\mathscr{J}_p$ or $I+\mathscr{J}_p$, $1<p<\infty$.*

(i) *$\mathscr{M}_p^\alpha$ is a closed operator on L^p for all $\alpha\in\mathbf{C}$.*

(ii) *If k is a positive integer, $\mathcal{M}_p^k$ is the k-th iterate of $\mathcal{M}_p$; that is,* Dom $(\mathcal{M}_p^k)$ *is defined inductively to be the set of all $f \in$* Dom $(\mathcal{M}_p^{k-1})$ *such that $\mathcal{M}_p^{k-1} f \in$* Dom $(\mathcal{M}_p)$, *and $\mathcal{M}_p^k f = \mathcal{M}_p \mathcal{M}_p^{k-1} f$.*

(iii) *If $f \in$* Dom $(\mathcal{M}_p^\beta) \cap$ Dom $(\mathcal{M}_p^{\alpha+\beta})$ *then $\mathcal{M}_p^\beta f \in$* Dom $(\mathcal{M}_p^\alpha)$ *and $\mathcal{M}_p^\alpha \mathcal{M}_p^\beta f = \mathcal{M}_p^{\alpha+\beta} f$. $\mathcal{M}_p^{\alpha+\beta}$ is the smallest closed extension of $\mathcal{M}_p^\alpha \mathcal{M}_p^\beta$. In particular, $\mathcal{M}_p^{-\alpha} = (\mathcal{M}_p^\alpha)^{-1}$.*

(iv) *If* $\operatorname{Re} \alpha < \operatorname{Re} \beta$ *and $f \in$* Dom $(\mathcal{M}_p^\alpha) \cap$ Dom $(\mathcal{M}_p^\beta)$, *then $f \in$* Dom $(\mathcal{M}_p^\gamma)$ *whenever* $\operatorname{Re} \alpha \leqq \operatorname{Re} \gamma \leqq \operatorname{Re} \beta$, *and $\|\mathcal{M}_p^\gamma f\|_p \leqq C \|\mathcal{M}_p^\alpha f\|_p^\theta \|\mathcal{M}_p^\beta f\|_p^{1-\theta}$ where $\theta = \operatorname{Re}(\gamma-\alpha)/\operatorname{Re}(\beta-\alpha)$ and C depends only on α, β,* $\arg(\gamma-\alpha)$, *and* $\arg(\beta-\gamma)$. *Moreover, $\mathcal{M}_p^\gamma f$ is an analytic L^p-valued function of γ on the strip* $\operatorname{Re} \alpha < \operatorname{Re} \gamma < \operatorname{Re} \beta$ *and is continous on the closure.*

(v) *$\mathcal{M}_p^\alpha$ is the (real) dual operator of $\mathcal{M}_{p'}^\alpha$, $p' = p/(p-1)$.*

(vi) *If $f \in$* Dom $(\mathcal{M}_p^\alpha) \cap L^q$ *then $f \in$* Dom $(\mathcal{M}_q^\alpha)$ *if and only if $\mathcal{M}_p^\alpha f \in L^q$, in which case $\mathcal{M}_p^\alpha f = \mathcal{M}_q^\alpha f$.*

Proof. In [15] Komatsu considers a closed operator A on a Banach space for which $(-\infty, 0)$ is in the resolvent set and $\|\lambda(\lambda I + A)^{-1}\|$ is bounded independently of λ for $\lambda > 0$. He defines closed fractional powers A_+^α ($\operatorname{Re} \alpha > 0$), A_-^α ($\operatorname{Re} \alpha < 0$), and A_0^α ($\alpha \in \mathbf{C}$); however, in case A has dense domain and range, $A_+^\alpha = A_0^\alpha$ for $\operatorname{Re} \alpha > 0$ and $A_-^\alpha = A_0^\alpha$ for $\operatorname{Re} \alpha < 0$ ([15], Proposition 4.12), and the subscripts may be dropped. We take $A = \mathcal{M}_p$, which has dense domain and satisfies $\|\lambda(\lambda I + \mathcal{M}_p)^{-1}\| \leqq$ const. for $\lambda > 0$ by the Hille—Yosida theorem [30], and which has dense range (by Proposition (2.19) and Theorem (3.8) for $\mathcal{M}_p = \mathcal{J}_p$, and because -1 is in the resolvent set of $\mathcal{J}_p$ for $\mathcal{M}_p = I + \mathcal{J}_p$). For the moment, we denote the α-th power of $\mathcal{M}_p$ as defined in [15] by $\tilde{\mathcal{M}}_p^\alpha$.

Since $\{H_t\}$ is an analytic semigroup on L^p, it follows from Proposition 4.12 of [15], Theorem 5.4 of [16], and Theorem 6.3 of [17] that $\mathcal{M}_p^\alpha = \tilde{\mathcal{M}}_p^\alpha$ for $\operatorname{Re} \alpha \neq 0$, and that (ii) is true. (i) is therefore also true (even for $\operatorname{Re} \alpha = 0$, since then $\mathcal{M}_p^\alpha$ is bounded). (v) is clear when $\operatorname{Re} \alpha = 0$, and it is true for $\alpha = 1$ by Phillips' theorem [30] since H_t is self-adjoint; the general case then follows from Theorem 2.10 of [19]. (vi) is true for $\alpha = 1$ by Theorem (3.8); moreover, the semigroups generated by $\mathcal{M}_p$ and $\mathcal{M}_q$ coincide on $L^p \cap L^q$, and therefore the resolvents $(\lambda I + \mathcal{M}_p)^{-1}$ and $(\lambda I + \mathcal{M}_q)^{-1}$ are equal on $L^p \cap L^q$ for $0 < \lambda < \infty$. (vi) then follows for $\operatorname{Re} \alpha \neq 0$ from Theorems 3.2 and 3.3 of [20], and it is obvious for $\operatorname{Re} \alpha = 0$.

(iii) is true with $\mathcal{M}_p^\alpha$ replaced by $\tilde{\mathcal{M}}_p^\alpha$ by Theorems 7.2 and 7.3 of [15]. Also, it follows from Theorems 8.1 and 8.2 of [15] that if $\operatorname{Re} \alpha < \operatorname{Re} \beta$ and $f \in$ Dom $(\tilde{\mathcal{M}}_p^\alpha) \cap$ $\cap$ Dom $(\tilde{\mathcal{M}}_p^\beta)$ then $f \in$ Dom $(\tilde{\mathcal{M}}_p^\gamma)$ for $\operatorname{Re} \alpha < \operatorname{Re} \gamma < \operatorname{Re} \beta$ and $\|\tilde{\mathcal{M}}_p^\gamma f\|_p \leqq C \|\tilde{\mathcal{M}}_p^\alpha f\|_p^\theta \|\tilde{\mathcal{M}}_p^\beta f\|_p^{1-\theta}$ with C, θ as in (iv); moreover, $\tilde{\mathcal{M}}_p^\gamma f$ is analytic in γ in this strip and $\tilde{\mathcal{M}}_p^\gamma f \to \tilde{\mathcal{M}}_p^\alpha f$ (or $\tilde{\mathcal{M}}_p^\gamma f \to \tilde{\mathcal{M}}_p^\beta f$) as $\gamma \to \alpha$ ($\gamma \to \beta$) in a region $|\arg(\gamma-\alpha)| \leqq$ $\leqq C < \pi/2$ ($|\arg(\gamma-\beta)| \leqq C < \pi/2$). If we show that $\mathcal{M}_p^\alpha = \tilde{\mathcal{M}}_p^\alpha$ for all α, then, (iv) will follow, since $\mathcal{M}_p^\alpha f$ is clearly continuous in α along the line $\operatorname{Re} \alpha = 0$.

It therefore remains to prove that $\mathcal{M}_p^\alpha=\tilde{\mathcal{M}}_p^\alpha$ for $\operatorname{Re}\alpha=0$, and for this it suffices to show that $\mathcal{M}_p^\alpha f=\tilde{\mathcal{M}}_p^\alpha f$ for all f in a dense subspace $V\subset L^p$. For $\mathcal{M}_p=\mathcal{J}_p$ we take $V=\mathcal{J}(\mathcal{D})$, which is dense by Proposition (2.19), and for $\mathcal{M}_p=I+\mathcal{J}_p$ we take $V=\mathcal{D}$. In either case, $V\subset\operatorname{Dom}(\mathcal{M}_p)\cap\operatorname{Dom}(\mathcal{M}_p^{-1})$ for all p; hence (since $\mathcal{M}_p^{-1}=\tilde{\mathcal{M}}_p^{-1}$) $V\subset\operatorname{Dom}(\tilde{\mathcal{M}}_p^\alpha)$ for $|\operatorname{Re}\alpha|<1$ and $1<p<\infty$. It is clear that $\mathcal{M}_2^\alpha f=\tilde{\mathcal{M}}_2^\alpha f$ for $\operatorname{Re}\alpha=0$ and $f\in V$, since both sides are analytic functions of α which agree for $\operatorname{Re}\alpha\neq0$. But then by the assertions proved above, if $f\in V$, $\operatorname{Re}\alpha=0$, and $0<\varepsilon<1$,

$$\tilde{\mathcal{M}}_p^\alpha f=\tilde{\mathcal{M}}_p^{\alpha+\varepsilon}\tilde{\mathcal{M}}_p^{-\varepsilon}f=\tilde{\mathcal{M}}_2^{\alpha+\varepsilon}\tilde{\mathcal{M}}_2^{-\varepsilon}f=\tilde{\mathcal{M}}_2^\alpha f=\mathcal{M}_2^\alpha f=\mathcal{M}_p^\alpha f.$$

This completes the proof.

(3.16) **Proposition.** *If* $\operatorname{Re}\alpha>0$ *and* $1<p<\infty$, $\operatorname{Dom}(\mathcal{J}_p^\alpha)=\operatorname{Dom}\big((I+\mathcal{J}_p)^\alpha\big)$.

Proof. This is an instance of Theorem 6.4 of Komatsu [15].

By Theorem (3.15. (vi)), $\mathcal{J}_p^\alpha$ agrees with $\mathcal{J}_q^\alpha$, and $(I+\mathcal{J}_p)^\alpha$ with $(I+\mathcal{J}_q)^\alpha$, on their common domains for $\alpha\in\mathbf{C}$ and $1<p,\ q<\infty$. We shall therefore omit the subscripts on these operators except when we wish to specify domains.

In certain cases we can express $\mathcal{J}^\alpha$ as an integral operator with homogeneous kernel, as is shown by the next two propositions.

(3.17) **Proposition.** *Suppose* $0<\operatorname{Re}\alpha<Q$. *The integral*

$$R_\alpha(x)=\frac{1}{\Gamma(\alpha/2)}\int_0^\infty t^{(\alpha/2)-1}h(x,t)\,dt$$

converges absolutely for all $x\neq0$, *and* R_α *is a kernel of type* α.

Proof. It follows from (3.2) that $h(x,t)=O(t^{-Q/2})$ as $t\to\infty$ for each x, and by Corollary (3.4), $h(x,t)=O(t^N)$ for all N as $t\to0$, for each $x\neq0$. Hence $t^{(\alpha/2)-1}h(x,t)$ is in L^1 as a function of t for each $x\neq0$ provided $\operatorname{Re}\alpha<Q$, so $R_\alpha(x)$ exists for $x\neq0$. Likewise, if D is a homogeneous differential operator of degree k on G, $Dh(x,t)=O(t^{-(Q+k)/2})$ as $t\to\infty$ and $Dh(x,t)=O(t^N)$ as $t\to0$ for $x\neq0$. The integral

$$\int_0^\infty t^{(\alpha/2)-1}Dh(x,t)\,dt$$

thus converges locally uniformly in x away from $x=0$, so we may differentiate under the integral sign and conclude that R_α is $\mathscr{C}^\infty$ on $G-\{0\}$. Finally, by (3.2),

$$R_\alpha(rx)=\frac{1}{\Gamma(\alpha/2)}\int_0^\infty t^{(\alpha/2)-1}h(rx,t)\,dt=\frac{1}{\Gamma(\alpha/2)}\int_0^\infty t^{(\alpha/2)-1}r^{-Q}h(x,r^{-2}t)\,dt$$
$$=\frac{1}{\Gamma(\alpha/2)}r^{\alpha-Q}\int_0^\infty s^{(\alpha/2)-1}h(x,s)\,ds=r^{\alpha-Q}R_\alpha(x).$$

Thus by Proposition (1.8), R_α is a kernel of type α, $0<\operatorname{Re}\alpha<Q$.

(3.18) **Proposition.** *Suppose* $f \in L^p$ $(1<p<\infty)$ *and the integral*

$$g(x) = f * R_\alpha(x) = \int f(xy^{-1}) R_\alpha(y)\, dy \quad (0 < \operatorname{Re}\alpha < Q)$$

converges absolutely for almost every x. *If* $f \in \operatorname{Dom}(\mathcal{J}_p^{-\alpha/2})$ *then* $g \in L^p$ *and* $\mathcal{J}^{-\alpha/2} f = g$.

Proof. Set $R_\alpha^\eta(x) = [\Gamma(\alpha/2)]^{-1} \int_0^\eta t^{(\alpha/2)-1} h(x, t)\, dt$. By Theorem (3.1.(i)), $\|R_\alpha^\eta\|_1 = [\Gamma(\alpha/2)]^{-1} \int_0^\eta t^{(\alpha/2)-1} dt < \infty$, so that if $f \in L^p$, $g_\eta = f * R_\alpha^\eta$ is defined almost everywhere and is in L^p. Also by Fubini's theorem we have

$$g_\eta = [\Gamma(\alpha/2)]^{-1} \int_0^\eta t^{(\alpha/2)-1} H_t f\, dt.$$

Suppose $f \in \operatorname{Dom}(\mathcal{J}_p^{-\alpha/2})$. Then $g_\eta \to \mathcal{J}^{-\alpha/2} f$ in the L^p norm as $\eta \to \infty$, so it will suffice to show that for every sequence $\eta_j \to \infty$, $g_{\eta_j} \to g$ almost everywhere. Given such a sequence, we can find a set E of measure zero such that for all $x \notin E$, the integrals defining $g(x)$ and $g_{\eta_j}(x)$ $(j=1, 2, 3, \dots)$ are all absolutely convergent. That is, the functions

$$\psi(y, t) = \frac{1}{\Gamma(\alpha/2)} t^{(\alpha/2)-1} f(xy^{-1}) h(y, t), \quad \psi_j(y, t) = \psi(y, t) \chi_j(t)$$

are in $L^1(G \times (0, \infty))$, where χ_j is the characteristic function of $[0, \eta_j]$. But $|\psi_j(y, t)| \leqq |\psi(y, t)|$, so by Lebesgue's theorem,

$$g_{\eta_j}(x) = \iint \psi_j(y, t)\, dy\, dt \to \iint \psi(y, t)\, dy\, dt = g(x),$$

and we are done.

The kernels R_α are a generalization of the classical Riesz potentials (cf. Stein [25]), which are obtained by taking $G = \mathbf{R}^n$ and h to be the usual heat kernel as in (2.11). The kernel R_2 is of course the fundamental solution for $\mathcal{J}$ given by Theorem (2.1).

We can also define generalized Bessel potentials $J_\alpha(x)$ for $\operatorname{Re}\alpha > 0$ by

$$J_\alpha(x) = \frac{1}{\Gamma(\alpha/2)} \int_0^\infty t^{(\alpha/2)-1} e^{-t} h(x, t)\, dt.$$

Arguments similar to the ones above then yield the following properties of J_α, although we shall not insist on the details:

(1) $J_\alpha(x)$ is defined for all $x \neq 0$, and even for $x=0$ in case $\operatorname{Re}\alpha > Q$. Moreover J_α is $\mathscr{C}^\infty$ away from 0.

(2) As $x \to 0$, $|J_\alpha(x)| = O(|x|^{\operatorname{Re}\alpha - Q})$ if $\operatorname{Re}\alpha < Q$, $|J_\alpha(x)| = O(\log(1/|x|))$ if $\operatorname{Re}\alpha = Q$, and J_α is continuous at 0 if $\operatorname{Re}\alpha > Q$. As $x \to \infty$, $|J_\alpha(x)| = O(|x|^{-N})$ for all N. (Hence $J_\alpha \in L^1$ for all α, $\operatorname{Re}\alpha > 0$.)

(3) If $f \in L^p$, $1 < p < \infty$, then $(I + \mathcal{J})^{-\alpha/2} f = f * J_\alpha$.

4. Sobolev spaces

In this section we develop a theory of potential spaces of Sobolev type in terms of the sub-Laplacian on the stratified group G.

Definition. For $1<p<\infty$ and $\alpha\geqq 0$, S_α^p is the space Dom $(\mathscr{J}_p^{\alpha/2})$ equipped with the graph norm $\|f\|_{p,\alpha}=\|f\|_p+\|\mathscr{J}^{\alpha/2}f\|_p$.

By Theorem (3.15. (i)), S_α^p is a Banach space. An alternative characterization of S_α^p which will often be convenient is the following.

(4.1) **Proposition.** *$S_\alpha^p=$ Dom $((I+\mathscr{J}_p)^{\alpha/2})$, and the norms $\|f\|_{p,\alpha}$ and $\|(I+\mathscr{J})^{\alpha/2}f\|_p$ are equivalent.*

Proof. The first assertion is just Proposition (3.16), and it then follows from the closed graph theorem that the norms $\|f\|_{p,\alpha}$ and $\|f\|_p+\|(I+\mathscr{J})^{\alpha/2}f\|_p$ are equivalent. But $\|f\|_p\leqq C\|(I+\mathscr{J})^{\alpha/2}f\|_p$ since $(I+\mathscr{J})^{-\alpha/2}$ is bounded.

Remark. This proposition, together with (iii) and (iv) of Theorem (3.15), shows that for all $\alpha,\beta\geqq 0$, $(I+\mathscr{J})^{\beta/2}$ is an isomorphism of $S_{\alpha+\beta}^p$ with S_α^p. We can therefore define S_α^p for $\alpha<0$ to be the completion of L^p with respect to the norm $\|(I+\mathscr{J})^{\alpha/2}f\|_p$. By Theorem (3.15. (v)), then, S_α^p is the dual space of $S_{-\alpha}^{p'}$ where $p'=p/(p-1)$. Theorems about S_α^p for $\alpha<0$ can thus be derived from those about S_α^p for $\alpha\geqq 0$ by duality arguments. In what follows we shall always assume $\alpha\geqq 0$ and leave the extensions to $\alpha<0$ to the reader.

We now derive some basic properties of S_α^p.

(4.2) **Proposition.** *If $0\leqq\gamma<\beta$ then $S_\beta^p\subset S_\gamma^p$ and $\|f\|_{p,\gamma}\leqq C_{p,\beta,\gamma}\|\|f\|_{p,\beta}$.*

Proof. Obvious from Theorem (3.15. (iv)), taking $\alpha=0$.

(4.3) **Proposition.** *If $a\leqq \operatorname{Re}\beta\leqq b\leqq 0$ then $(I+\mathscr{J})^\beta$ is bounded on S_α^p for all p, α with bound $\leqq C|\Gamma(1-i\operatorname{Im}\beta)|^{-1}$ where C depends only on p, α, a, and b.*

Proof. By Proposition (4.1), boundedness of $(I+\mathscr{J})^\beta$ on S_α^p is equivalent with boundedness of $(I+\mathscr{J})^{\alpha/2}(I+\mathscr{J})^\beta(I+\mathscr{J})^{-\alpha/2}=(I+\mathscr{J})^\beta$ on L^p. Moreover, $\|(I+\mathscr{J})^\beta\|\leqq$ $\leqq\|(I+\mathscr{J})^{\operatorname{Re}\beta}\|\,\|(I+\mathscr{J})^{i(\operatorname{Im}\beta)}\|$, so the assertion follows from Proposition (3.14) and the smooth dependence of $(I+\mathscr{J})^{\operatorname{Re}\beta}$ on $\operatorname{Re}\beta$.

(4.4) **Proposition.** *If $f\in L^p$ then $H_tf\in S_\beta^p$ for all $\beta\geqq 0$, $t>0$. If also $f\in S_\alpha^p$ then $H_tf\to f$ in the S_α^p norm as $t\to 0$.*

Proof. Suppose $f\in L^p$. Since $\mathscr{J}^kH_tf=f*\mathscr{J}^kh_t$ and $\mathscr{J}^kh_t\in L^1$ by Corollary (3.6), we have $\mathscr{J}^kH_tf\in L^p$ for $k=1,2,3,\dots$, hence (by Theorems (3.8) and (3.15. (ii))

$H_t f \in S^p_{2k}$ for all k. But then by Proposition (4.2), $H_t f \in S^p_\beta$ for all $\beta \geqq 0$. If $f \in S^p_\alpha$ and $k = [\alpha/2]+1$,

$$H_t \mathcal{J}^{\alpha/2} f = H_t \left[\lim_{\varepsilon \to 0} \frac{1}{\Gamma(k-(\alpha/2))} \int_\varepsilon^\infty s^{k-(\alpha/2)-1} \mathcal{J}^k H_s f \, ds \right]$$

$$= \lim_{\varepsilon \to 0} \frac{1}{\Gamma(k-(\alpha/2))} \int_\varepsilon^\infty s^{k-(\alpha/2)-1} \mathcal{J}^k H_{s+t} f \, ds = \mathcal{J}^{\alpha/2} H_t f$$

since H_t commutes with $\mathcal{J}^k$ on Dom $(\mathcal{J}^k_p)$. Thus as $t \to 0$, $H_t f \to f$ and $\mathcal{J}^{\alpha/2} H_t f = H_t \mathcal{J}^{\alpha/2} f \to \mathcal{J}^{\alpha/2} f$ in the L^p norm, that is, $H_t f \to f$ in S^p_α.

(4.5) **Theorem.** *$\mathcal{D}$ is a dense subspace of S^p_α for all p, α.*

Proof. As in the preceding proof we have $\mathcal{D} \subset$ Dom $(\mathcal{J}^k_p)$ for $k=1, 2, 3, \ldots$ and hence $\mathcal{D} \subset S^p_\alpha$ for all p, α. By Propositions (4.2) and (4.4), in order to show that $\mathcal{D}$ is dense in S^p_α it will suffice to show that $\mathcal{D}$ is dense in $\{H_t f : f \in S^p_\alpha, t>0\}$ in the S^p_{2k} norm, where k is an integer larger than $\alpha/2$.

If $f \in S^p_\alpha$ and $t>0$ then $H_t f$ is a solution of the hypoelliptic equation $(\mathcal{J}+(\partial/\partial t))H_t f = 0$ and hence is $\mathcal{C}^\infty$. Also, all left-invariant derivatives of $H_t f$ are in L^p, being convolutions of f with L^1 functions by Corollary (3.6). Choose $\varphi \in \mathcal{D}$ with $\varphi = 1$ on a neighborhood of 0, and set $g_\varepsilon(x) = \varphi(\varepsilon x) H_t f(x)$ for $\varepsilon > 0$. Then $g_\varepsilon \in \mathcal{D}$ and $g_\varepsilon \to H_t f$ in L^p as $\varepsilon \to 0$. Also,

$$\mathcal{J}^k g_\varepsilon(x) = \varphi(\varepsilon x)(\mathcal{J}^k H_t f)(x) + \sum_{j=2}^{2k} \varepsilon^j (D^j \varphi)(\varepsilon x)(D^{2k-j} H_t f)(x)$$

where D^i is a homogeneous operator of degree i, $0 \leqq i \leqq 2k$. The first term tends to $\mathcal{J}^k H_t f$ in L^p as $\varepsilon \to 0$, and the other terms tend to zero because of the factors of ε^j. Thus $g_\varepsilon \to H_t f$ in S^p_{2k}, and we are done.

(4.6) **Corollary.** *If $f \in L^p$ and $g \in \mathcal{D}$ then $f * g \in S^p_\alpha$ for all α.*

Proof. Choose a sequence $\{f_j\} \subset \mathcal{D}$ with $f_j \to f$ in L^p. Then $f_j * g \in \mathcal{D}$, $f_j * g \to f * g$ in L^p, and $\mathcal{J}^k(f_j * g) = f_j * \mathcal{J}^k g \to f * \mathcal{J}^k g$ in L^p. Hence $f * g \in S^p_{2k}$ for all k, so $f * g \in S^p_\alpha$ for all α.

We now prove the fundamental interpolation theorem for operators on the S^p_α spaces. This result is due to Calderón [2] in the Euclidean case, and our proof is an adaptation of his.

(4.7) **Theorem.** *Let G_1 and G_2 be stratified groups with sub-Laplacians $\mathcal{J}_{(1)}$ and $\mathcal{J}_{(2)}$. Let T be a linear mapping from $S^{p_0}_{\alpha_0}(G_1) + S^{p_1}_{\alpha_1}(G_1)$ to locally integrable functions on G_2, and suppose T maps $S^{p_0}_{\alpha_0}(G_1)$ and $S^{p_1}_{\alpha_1}(G_1)$ boundedly into $S^{q_0}_{\beta_0}(G_2)$ and $S^{q_1}_{\beta_1}(G_2)$, respectively, Then T extends uniquely to a bounded mapping from $S^{p_t}_{\alpha_t}(G_1)$ to $S^{q_t}_{\beta_t}(G_2)$ for $0 \leqq t \leqq 1$, where*

$$(\alpha_t, \beta_t, p_t^{-1}, q_t^{-1}) = t(\alpha_1, \beta_1, p_1^{-1}, q_1^{-1}) + (1-t)(\alpha_0, \beta_0, p_0^{-1}, q_0^{-1}) \in \mathbf{R}^4.$$

Proof. Let $B=L^1(G_1)\cap L^\infty(G_1)$, which is a dense subspace of $L^p(G_1)$ for $p<\infty$ and includes all step functions. Choose $\varphi\in\mathscr{D}$ with $\int\varphi(x)\,dx=1$, and set $\varphi_\varepsilon(x)=\varepsilon^{-Q}\varphi(\varepsilon^{-1}x)$ for $\varepsilon>0$; by a standard argument, $\{\varphi_\varepsilon\}$ is an approximation to the identity as $\varepsilon\to 0$. For each $\varepsilon>0$ we define the family $\{T_z^\varepsilon: 0\leqq \operatorname{Re} z\leqq 1\}$ of operators on B by

$$T_z^\varepsilon f=(I+\mathscr{J}_{(2)})^{\beta_z}T(I+\mathscr{J}_{(1)})^{-\alpha_z}(f*\varphi_\varepsilon)$$

where $2\alpha_z=z\alpha_1+(1-z)\alpha_0$ and $2\beta_z=z\beta_1+(1-z)\beta_0$. T_z^ε is well defined on B by Proposition (4.3) and Corollary (4.6). For $0\leqq\operatorname{Re} z\leqq 1$ let us set

$$\begin{aligned}A(z)&=\left|\Gamma(1+i(\operatorname{Im}\alpha_z))\Gamma(1-i(\operatorname{Im}\beta_z))\right|^{-1}\\&=\left|\Gamma\left(1+\frac{1}{2}i(\operatorname{Im} z)(\alpha_1-\alpha_0)\right)\Gamma\left(1-\frac{1}{2}i(\operatorname{Im} z)(\beta_1-\beta_0)\right)\right|^{-1}.\end{aligned}$$

Then, supposing for the sake of definiteness that $\beta_1\geqq\beta_0$, Proposition (4.3) implies that $T_z^\varepsilon f\in L^{q_1}(G_2)$ for $f\in B$ and

$$\|T_z^\varepsilon f\|_{q_1}\leqq CA(z)\|f*\varphi_\varepsilon\|_{p_1,\alpha_1}\quad (C \text{ independent of } f \text{ and } z).$$

Thus by Theorem (3.15. (iv)), for any $f\in B$ and $g\in L^{q_1'}(G_2)$ (where $q_1'=q_1/(q_1-1)$) the mapping $z\to\int_{G_2}(T_z^\varepsilon f)g$ is analytic for $0<\operatorname{Re} z<1$ and continuous for $0\leqq\operatorname{Re} z\leqq 1$ and satisfies

$$\left|\int_{G_2}(T_z^\varepsilon f)g\right|\leqq C\|f*\varphi_\varepsilon\|_{p_1,\alpha_1}\|g\|_{q_1'}A(z).$$

Moreover, by the hypotheses of the theorem and Proposition (4.3), for any $s\in\mathbf{R}$, $T_{is}^\varepsilon f\in L^{q_0}(G_2)$ and $T_{1+is}^\varepsilon f\in L^{q_1}(G_2)$ for all $f\in B$, and we have the estimates

$$\|T_{is}^\varepsilon f\|_{q_0}\leqq C_0A(is)\|f*\varphi_\varepsilon\|_{p_0}\leqq C_0A(is)\|f\|_{p_0},$$

$$\|T_{1+is}^\varepsilon f\|_{q_1}\leqq C_1A(1+is)\|f*\varphi_\varepsilon\|_{p_1}\leqq C_1A(1+is)\|f\|_{p_1},$$

where C_0 and C_1 are independent of f, s, and ε. Therefore, since $A(z)=O(e^{\pi|\operatorname{Im} z|})$ by Stirling's formula, the Riesz—Thorin—Stein interpolation theorem (cf. [28] or [31]) implies that for $0\leqq t\leqq 1$, $T_t^\varepsilon f\in L^{q_t}(G_2)$ for all $f\in B$ and

$$(4.8)\qquad \|T_t^\varepsilon f\|_{q_t}\leqq C_t\|f\|_{p_t}$$

where C_t depends only on t, C_0, C_1, and the function A.

Now, for $1<p<\infty$, let

$$\mathscr{V}_p=\{g=f*\varphi_\varepsilon: f\in B,\ \varepsilon>0, \text{ and } \|f\|_p\leqq 2\|f*\varphi_\varepsilon\|_p\}.$$

We note that for any $f\in B$ and $1<p<\infty$, $f*\varphi_\varepsilon\in\mathscr{V}_p$ for ε sufficiently small, since $f*\varphi_\varepsilon\to f$ in L^p. In particular, $\mathscr{V}_p$ is dense in L^p for all p. (4.8) then says that if $g=f*\varphi_\varepsilon\in\mathscr{V}_{p_t}$, $0\leqq t\leqq 1$,

$$\|(I+\mathscr{J}_{(2)})^{\beta_t}T(I+\mathscr{J}_{(1)})^{-\alpha_t}g\|_{p_t}\leqq C_t\|f\|_{p_t}\leqq 2C_t\|g\|_{p_t}.$$

Hence $(I+\mathcal{J}_{(2)})^{\beta_t}T(I+\mathcal{J}_{(1)})^{-\alpha_t}$ extends uniquely to a bounded mapping from $L^{p_t}(G_1)$ to $L^{q_t}(G_2)$. But by Proposition (4.1), this means that T extends uniquely to a bounded mapping from $S^{p_t}_{\alpha_t}(G_1)$ to $S^{q_t}_{\beta_t}(G_2)$, and the proof is complete.

We return to the case of a single stratified group G. Let K be a kernel of type 0; we recall (Proposition (1.9)) that the mapping $T_K: f \to f*K$ is bounded on L^p, $1<p<\infty$. Our next objective is to extend this result to S^p_α for all $\alpha \geqq 0$. This is easy in the Abelian case, since then T_K commutes with constant-coefficient differential operators, but the general situation requires a more substantial argument.

(4.9) **Theorem.** *Let K be any kernel of type 0, and let $T_K: f \to f*K$ be the associated operator on L^p, $1<p<\infty$. If $f\in S^p_\alpha$ $(\alpha \geqq 0)$ then $T_K f \in S^p_\alpha$ and $\|T_K f\|_{p,\alpha} \leqq C_{p,\alpha}\|f\|_{p,\alpha}$.*

Proof. It suffices to prove that for any kernel K of type 0, T_K is a bounded operator on S^p_α for $\alpha=0, 2, 4, 6, \ldots$ and $1<p<\infty$, as Theorem (4.7) then implies the general result. Moreover, by Theorem (4.5) it is enough to show that if $u\in\mathcal{D}$ then $T_K u\in S^p_\alpha$ for $\alpha=0, 2, 4, 6, \ldots$ and $1<p<\infty$, and that $\|T_K u\|_{p,\alpha} \leqq C_{p,\alpha}\|u\|_{p,\alpha}$.

We proceed by induction, the initial step $\alpha=0$ being Proposition (1.9). Assume then that the theorem is proved for $\alpha=0, 2, 4, \ldots, 2j$, and suppose K is a kernel of type 0 and $u\in\mathcal{D}$. Then $\mathcal{J}u\in S^p_{2j}\cap \mathrm{Dom}\,(\mathcal{J}_p^{-1})$, and by Proposition (1.11), $\mathcal{J}u*R_2$ exists a.e. Thus by Proposition (3.18), $u=\mathcal{J}^{-1}\mathcal{J}u=\mathcal{J}u*R_2$, and then by Proposition (1.13), $T_K u=(\mathcal{J}u*R_2)*K=\mathcal{J}u*(R_2*K)$. Now R_2*K is a kernel of type 2, so $K_0=\mathcal{J}(R_2*K)$ is a kernel of type 0, and we have

$$\mathcal{J}T_K u = T_K u*\mathcal{J}\delta = \mathcal{J}u*\big((R_2*K)*\mathcal{J}\delta\big) = \mathcal{J}u*K_0 = T_{K_0}(\mathcal{J}u),$$

since $\mathcal{J}u$ and $\mathcal{J}\delta$ have compact support. By inductive hypothesis, then, $T_K u\in S^p_{2j}$, and we already know $T_K u\in L^p$. Therefore $T_K u\in S^p_{2j+2}$ and

$$\begin{aligned}\|T_K u\|_{p,2j+2} &\leqq C(\|T_K u\|_p+\|\mathcal{J}T_K u\|_{p,2j}) \leqq C(\|T_K u\|_p+\|T_{K_0}\mathcal{J}u\|_{p,2j})\\ &\leqq C'(\|u\|_p+\|\mathcal{J}u\|_{p,2j}) \leqq C''\|u\|_{p,2j+2}.\end{aligned}$$

The proof is complete.

The next theorem provides a characterization of S^p_α in terms of left-invariant derivatives. We recall that $\mathcal{J}=-\sum_1^n X_j^2$ where $X_1, \ldots, X_n$ is a basis for V_1, and that $\tilde{X}$ denotes the right-invariant vector field agreeing with the left-invariant vector field X at 0.

(4.10) **Theorem.** *If $1<p<\infty$ and $\alpha\geqq 0$, then $f\in S^p_{\alpha+1}$ if and only if f and the distribution derivatives Xf are in S^p_α for all $X\in V_1$. The norms $\|f\|_{p,\alpha+1}$ and $\|f\|_{p,\alpha}+$ $+\sum_1^n \|X_j f\|_{p,\alpha}$ are equivalent.*

Before proceeding to the proof, we need two technical lemmas.

(4.11) **Lemma.** *If $u\in\mathcal{D}$ and $0<\alpha\leqq 2$ then $\mathcal{J}^{\alpha/2}u=\mathcal{J}(u*R_{2-\alpha})$.*

Proof. The proof of Proposition (2.19) shows that if $u\in\mathscr{D}$ there is a sequence $\{u_k\}\subset\mathscr{D}$ such that $u_k\to u*R_2$ pointwise, $|u_k(x)|=O(|x|^{2-Q})$ uniformly in k, and $\mathscr{J}u_k\to u$ in L^p, $1<p<\infty$. The first two conditions imply that $u_k\to u*R_2$ in L^p for $p>Q/(Q-2)$; hence $u\in\text{Range}\,(\mathscr{J}_p)=\text{Dom}\,(\mathscr{J}_p^{-1})$ for $p>Q/(Q-2)$. By Theorem (3.15. (iv)), then, $u\in\text{Dom}\,(\mathscr{J}_p^{(\alpha/2)-1})$ for $p>Q/(Q-2)$ and $0\leqq\alpha\leqq 2$, and by Proposition (3.18), $\mathscr{J}^{(\alpha/2)-1}u=u*R_{2-\alpha}$. But then $\mathscr{J}^{\alpha/2}u=\mathscr{J}\mathscr{J}^{(\alpha/2)-1}u=\mathscr{J}(u*R_{2-\alpha})$.

(4.12) **Lemma.** *There exist kernels $K_1, \dots, K_n$ of type* 1 *such that for all $\tau\in\mathscr{E}'$, $\tau=\sum_1^n (X_j\tau)*K_j$.*

Proof. Since R_2 is a fundamental solution for $\mathscr{J}$, by Corollary (2.8) we have

$$\tau = \tau*\mathscr{J}\delta*R_2 = -\sum_1^n \tau*X_j\delta*X_j\delta*R_2 = -\sum_1^n (X_j\tau)*(\tilde{X}_jR_2).$$

Thus we take $K_j=-\tilde{X}_jR_2$.

Proof of Theorem (4.10). By Theorem (4.5), it suffices to show that the norms $\|u\|_{p,\alpha+1}$ and $\|u\|_{p,\alpha}+\sum_1^n \|X_ju\|_{p,\alpha}$ are equivalent for $u\in\mathscr{D}$. First, if $u\in\mathscr{D}$ then $\mathscr{J}^{1/2}u\in\text{Dom}\,(\mathscr{J}_p^{-1/2})$ for all p, in particular for $p<Q/(Q-1)$, so by Propositions (1.11) and (3.18), $u=\mathscr{J}^{-1/2}\mathscr{J}^{1/2}u=\mathscr{J}^{1/2}u*R_1$. Then for $X\in V_1$,

$$Xu = (\mathscr{J}^{1/2}u*R_1)*X\delta = \mathscr{J}^{1/2}u*(R_1*X\delta) = \mathscr{J}^{1/2}u*XR_1,$$

the associativity being justified by approximating $\mathscr{J}^{1/2}u$ in L^p $(p<Q/(Q-1))$ by elements of $\mathscr{D}$. But XR_1 is a kernel of type 0, so by Theorem (4.9) and Proposition (4.2),

$$\|Xu\|_{p,\alpha} \leqq C_{p,\alpha}\|\mathscr{J}^{1/2}u\|_{p,\alpha} \leqq C_{p,\alpha}(\|\mathscr{J}^{1/2}u\|_p+\|\mathscr{J}^{(\alpha+1)/2}u\|_p) \leqq C'_{p,\alpha}\|u\|_{p,\alpha+1}.$$

On the other hand, by Proposition (1.13) and Lemmas (4.11) and (4.12),

$$\mathscr{J}^{1/2}u = \mathscr{J}(u*R_1) = -\mathscr{J}\sum_1^n (X_ju)*(K_j*R_1) = -\sum_1^n (X_ju)*\mathscr{J}(K_j*R_1).$$

Now K_j*R_1 is a kernel of type 2, so $\mathscr{J}(K_j*R_1)$ is a kernel of type 0. Thus by Theorem (4.9),

$$\|u\|_{p,\alpha+1} \leqq \|u\|_{p,\alpha}+\|\mathscr{J}^{1/2}u\|_{p,\alpha} \leqq \|u\|_{p,\alpha}+C_{p,\alpha}\sum_1^n \|X_ju\|_{p,\alpha},$$

and we are done.

This theorem has several important corollaries. Before stating them we need to introduce a multi-index notation for non-commuting derivatives. Namely, $I=(i_1, \dots, i_k)$ will denote a k-tuple with k arbitrary and $1\leqq i_j\leqq n$ for $j=1, \dots, k$, and we set $|I|=k$. We then define X_I to be $X_{i_1}X_{i_2}\dots X_{i_k}$ (where $X_1, \dots, X_n$ is the chosen basis for V_1), which is a homogeneous differential operator of degree $|I|$. We note that every left-invariant differential operator is a linear combination of X_I's and in particular that V_k is in the span of the X_I's with $|I|=k$.

(4.13). **Corollary.** *If k is a positive integer,*

$$S_k^p=\{f\in L^p\colon X_I f\in L^p \text{ for } |I|\le k\},$$

and the norms $\|f\|_{p,k}+\sum_{|I|\le k}\|X_I f\|_p$ are equivalent.

(4.14) **Corollary.** *S_α^p is independent of the choice of sub-Laplacian, for $1<p<\infty$ and $\alpha\ge 0$.*

Proof. Let $X_1,\dots,X_n$ and $Y_1,\dots,Y_n$ be two bases for V_1, and let $\mathcal{J}_{(1)}=-\sum_1^n X_j^2$, $\mathcal{J}_{(2)}=-\sum_1^n Y_j^2$. It is evident from Corollary (4.13) that if α is an integer, the identity mapping $S_\alpha^p(\mathcal{J}_{(1)})\to S_\alpha^p(\mathcal{J}_{(2)})$ is an isomorphism. The general case now follows from Theorem (4.7).

(4.15) **Corollary.** *If $\varphi\in\mathcal{D}$, multiplication by φ is a bounded operator on S_α^p for all p, α.*

Proof. This is clear by Corollary (4.13) if α is an integer and Theorem (4.7) then yields the general case.

If $U\subset G$ is an open set, $1<p<\infty$, and $\alpha\ge 0$, we define

$$S_\alpha^p(U,\mathrm{loc})=\{f\in\mathcal{D}'(U)\colon \varphi f\in S_\alpha^p \text{ for all } \varphi\in\mathcal{D}(U)\}.$$

We abbreviate $S_\alpha^p(G,\mathrm{loc})$ as $S_\alpha^p(\mathrm{loc})$. Corollary (4.15) says that S_α^p is localizable, i.e. that $S_\alpha^p\subset S_\alpha^p(U,\mathrm{loc})$ for all $U\subset G$.

We now compare the spaces S_α^p with the classical Sobolev spaces. We consider the Lie algebra $\mathfrak{g}$ as an Abelian Lie group under addition, fix a linear coordinate system $\{x_j\}$ on $\mathfrak{g}$, and set $\Delta=-\sum_1^n(\partial/\partial x_j)^2$. The spaces $L_\alpha^p=S_\alpha^p(\Delta)$ are then the standard L^p Sobolev spaces on $\mathfrak{g}$, cf. Stein [25]. Since $\exp\colon\mathfrak{g}\to G$ is a diffeomorphism, we can also regard L_α^p as a space of functions on G.

(4.16) **Theorem.** *$L_\alpha^p(\mathrm{loc})\subset S_\alpha^p(\mathrm{loc})\subset L_{\alpha/m}^p(\mathrm{loc})$ for $1<p<\infty$* and *$\alpha\ge 0$, where m is the number of steps in the stratification of G.*

Proof. It suffices to show that for any $\varphi\in\mathcal{D}$, $T_\varphi\colon f\to\varphi f$ is bounded from L_α^p to S_α^p and from S_α^p to $L_{\alpha/m}^p$. Any constant-coefficient differential operator of order k can be expressed as a linear combination of X_I's with $|I|\le mk$ and smooth coefficients, and conversely any X_I with $|I|\le k$ is a linear combination of constant-coefficient operators of order $\leqq k$ with smooth coefficients. In particular, the coefficients are bounded on supp φ, so it follows easily from Corollary (4.13) that $T_\varphi\colon L^p_\alpha \to S^p_\alpha$ when α is an integer, and $T_\varphi\colon S^p_\alpha \to L^p_{\alpha/m}$ when α/m is an integer. The proof is concluded by applying Theorem (4.7) with $G_1=\mathfrak{g}$ and $G_2=G$ or vice versa.

Several remarks are in order concerning this theorem:

(1) Theorem (4.16) is essentially local in character. The coefficients of the vector fields X_j with respect to the coordinates $\{x_i \circ \exp^{-1}\}$ are in fact polynomials, hence unbounded at infinity, so we have no control over integrability conditions at infinity.

(2) It is easy to convince oneself that Theorem (4.16) cannot be improved. For example, it is clear that $L^p_\alpha(\mathrm{loc}) \not\subset S^p_\beta(\mathrm{loc})$ when $\beta<\alpha$ and α is an integer. On the other hand, given a positive integer k and $1<p<\infty$, choose $\varphi\in\mathcal{D}$ with $\varphi=1$ on a neighborhood of 0 and set $f(x)=\varphi(x)|x|^{2mk-(Q/p)}$. Then $\mathcal{J}^{mk}f$ is homogeneous of degree $-Q/p$ near 0, smooth away from 0, and compactly supported, hence (by Corollary (1.7)) in L^q for $q<p$. Thus $f\in S^q_{2mk}$ for $q<p$. Also, if $\beta>0$ is small and $\alpha=2mk-\beta$, by Proposition (1.11) and (3.18) we have

$$\mathcal{J}^{\alpha/2}f = \mathcal{J}^{mk-(\beta/2)}f = (\mathcal{J}^{mk}f) * R_\beta .$$

It is then easy to check that $\mathcal{J}^{\alpha/2}f$ is $O(|x|^{\beta-(Q/p)})$ near 0, $O(|x|^{-Q+\beta})$ near ∞, and smooth in between, so that (for β small), $\mathcal{J}^{\alpha/2}f\in L^p$. Since $f\in L^p$, we have $f\in S^p_\alpha$ for all $\alpha<2mk$. But if $Y\in V_m$ then $Y^{2k}f$ is homogeneous of degree $-Q/p$ near 0 and does not vanish identically there (by homogeneity considerations, $Y^{2k-1}f$ cannot be constant along trajectories of Y). Thus $Y^{2k}f\notin L^p$, so $f\notin L^p_{2k}$.

(3) In the case $p=2$ and $\alpha=1$, weaker versions of Theorem (4.16) — which, however, are valid in more general situations — have been obtained by Hörmander [10], Kohn (see [5], Theorem 5.4.7), and Radkevič (see [22]). We conjecture that our sharper result should also be valid more generally.

We conclude this section with a theorem related to the classical fractional integration theorems of Hardy—Littlewood and Sobolev (see [25] and [31]).

(4.17) **Theorem.** $S^p_\alpha \subset S^q_\beta$, *and* $\|\ \|_{q,\beta} \leqq C\|\ \|_{p,\alpha}$ *for some* $C=C(p,q,\alpha)>0$ *provided* $1<p<q<\infty$ *and* $\beta=\alpha-Q(p^{-1}-q^{-1})\geqq 0$.

Proof. Suppose $f\in S^p_\alpha$. Then $\mathcal{J}^{(\alpha-\beta)/2}f\in S^p_\beta\subset L^p$ since $\alpha>\beta$, also $\mathcal{J}^{(\alpha-\beta)/2}f\in \mathrm{Dom}\,(\mathcal{J}_p^{(\beta-\alpha)/2})$. By Propositions (1.11) and (3.18) we see that $f=(\mathcal{J}^{(\alpha-\beta)/2}f)*R_{\alpha-\beta}\in L^q$, and $\|f\|_q\leqq C\|\mathcal{J}^{(\alpha-\beta)/2}f\|_p\leqq C\|f\|_{p,\alpha}$. Likewise, $\mathcal{J}^{\alpha/2}f\in L^p\cap \mathrm{Dom}\,(\mathcal{J}_p^{(\beta-\alpha)/2})$, so $\mathcal{J}^{\beta/2}f=(\mathcal{J}^{\alpha/2}f)*R_{\alpha-\beta}\in L^q$ and $\|\mathcal{J}^{\beta/2}f\|_q\leqq$ $\leqq C\|\mathcal{J}^{\alpha/2}f\|_p\leqq C\|f\|_{p,\alpha}$. By Theorem (3.15. (vi)), then, $f\in\mathrm{Dom}\,(\mathcal{J}_q^{\beta/2})=S^q_\beta$, and $\|f\|_{q,\beta}\leqq C\|f\|_{p,\alpha}$.

5. Lipschitz spaces

We recall the definition of the classical Lipschitz spaces Λ_α ($\alpha>0$) on the stratified group G, cf. Stein [25]. Here we identify G with the Euclidean space $\mathfrak{g}$ with Euclidean norm $\| \ \|$ and linear coordinates $\{x_j\}$ via the exponential map.

Let $\mathscr{BC}$ be the space of bounded continuous functions on G. For $0<\alpha<1$ we define

$$\Lambda_\alpha = \{f\in\mathscr{BC}: \sup_{x,y} |f(x+y)-f(x)|/\|y\|^\alpha < \infty\}.$$

For $\alpha=1$,

$$\Lambda_1 = \{f\in\mathscr{BC}: \sup_{x,y} |f(x+y)+f(x-y)-2f(x)|/\|y\| < \infty\}.$$

Finally, if $\alpha=k+\alpha'$ where $k=1, 2, 3, \ldots$ and $0<\alpha'\leqq 1$,

$$\Lambda_\alpha = \{f\in\Lambda_{\alpha'}: \partial^j f/\partial x_{i_1} \ldots \partial x_{i_j} \in \Lambda_{\alpha'} \text{ whenever } j\leqq k\}.$$

For our purposes it is better to use a different family of Lipschitz spaces which are more closely related to the homogeneous structure on G, following the ideas in Korányi—Vági [21], Stein [27], and Folland—Stein [7]. (Most of the results in this section are proved in [7] for the case where G is a Heisenberg group.) Here we use the group structure and homogeneous norm on G. For $0<\alpha<1$ we define

$$\Gamma_\alpha = \{f\in\mathscr{BC}: |f|_\alpha = \sup_{x,y} |f(xy)-f(x)|/|y|^\alpha < \infty\}.$$

For $\alpha=1$,

$$\Gamma_1 = \{f\in\mathscr{BC}: |f|_1 = \sup_{x,y} |f(xy)+f(xy^{-1})-2f(x)|/|y| < \infty\}.$$

Then Γ_α, $0<\alpha\leqq 1$, is a Banach space with norm $\|f\|_{\Gamma_\alpha}=\|f\|_\infty+|f|_\alpha$. If $\alpha=k+\alpha'$ with $k=1, 2, 3, \ldots$ and $0<\alpha'\leqq 1$,

$$\Gamma_\alpha = \{f\in\Gamma_{\alpha'}: X_I f\in\Gamma_{\alpha'} \text{ whenever } |I|\leqq k\}.$$

(Here we are using the notation for derivatives introduced after Theorem (4.10).) Γ_α is a Banach space with norm $\|f\|_{\Gamma_\alpha}=\|f\|_{\Gamma_{\alpha'}}+\sum_{|I|\leqq k}\|X_I f\|_{\Gamma_{\alpha'}}$. For $f\in\Gamma_\alpha$ we also set $|f|_\alpha=|f|_{\alpha'}+\sum_{|I|\leqq k}|X_I f|_{\alpha'}$.

To study the spaces Γ_α we need to draw some consequences from the Campbell—Hausdorff formula (cf. Hochschild [9]).

(5.1) **Lemma.** *There is a constant $A>0$ and an integer N such that any $x\in G$ can be written $x=\prod_1^N x_j$ with $x_j\in\exp(V_1)$ and $|x_j|\leqq A|x|$, $j=1, \ldots, N$.*

Proof. Let $B=\{Y\in V_1: |\exp Y|\leqq 1\}$. In terms of the stratification $\mathfrak{g}=\bigoplus_1^m V_j$ and the basis $X_1, \ldots, X_n$ for V_1 we define maps $\varphi^0, \varphi^1_{i_1}, \varphi^2_{i_1 i_2}, \ldots, \varphi^{m-1}_{i_1\ldots i_{m-1}}$ $(1\leqq i_j\leqq n)$ of B into G by

$$\varphi^0(Y) = \exp Y,$$

$$\varphi^j_{i_1\ldots i_j}(Y) = [\ldots[[\exp Y, \exp X_{i_1}], \exp X_{i_2}], \ldots \exp X_{i_j}],$$

where $[x, y]=xyx^{-1}y^{-1}$. By the Campbell—Hausdorff formula, the differential $D_0\varphi^j_{i_1\dots i_j}\colon V_1\to\mathfrak{g}$ of $\varphi^j_{i_1\dots i_j}$ at the origin is given by

$$D_0\varphi^0(Y) = Y,$$

$$D_0\varphi^j_{i_1\dots i_j}(Y) = [\dots[[Y, X_{i_1}], X_{i_2}], \dots X_{i_j}].$$

Now consider the map

$$\varphi\colon Y^j_{i_1\dots i_j} \to \prod_{j=0}^{m-1} \prod_{1\le i_k\le n,\ 1\le k\le j} \varphi^j_{i_1\dots i_j}(Y^j_{i_1\dots i_j})$$

from the $(\sum_0^{m-1} n^j)$-fold product of B with itself into G. The preceding remarks, together with another application of Campbell—Hausdorff, show that the differential $D_0\varphi$ is surjective onto $\mathfrak{g}$. Consequently, the range of φ includes a ball $\{x\colon |x|\le r_0\}$ of positive radius about the origin in G. Since a commutator of $j+1$ elements of G is the product of $3\cdot 2^j-2$ elements, any $x\in G$ with $|x|\le r_0$ can be written as the product of $N=\sum_{j=0}^{m-1} n^j(3\cdot 2^j-2)$ elements of $\exp(V_1)$ whose norms are at most 1. By dilation, then, an arbitrary x can be written as the product of N elements of $\exp(V_1)$ whose norms are at most $r_0^{-1}|x|$.

A similar (but easier) argument yields:

(5.2) **Lemma.** *For $1\le k\le m$, the mapping $(y_k, \dots, y_m)\to\prod_k^m y_j$ is a diffeomorphism from $(\exp V_k)\times\dots\times(\exp V_m)$ onto $\exp(\bigoplus_k^m V_j)$, and there is a constant $A>0$ such that if $y=\prod_k^m y_j$ with $y_j\in\exp V_j$ then $|y_j|\le A|y|$.*

If $x\in G$, we can write x uniquely as $x=\exp(X+Y)$ where $X\in V_1$ and $Y\in\bigoplus_2^m V_j$. We define the "partial inverse of x with respect to V_1", denoted $\tilde{x}$, to be $\tilde{x}=\exp(-X+Y)$.

(5.3) **Lemma.** *For any $x\in G$, $x\tilde{x}$ and $\tilde{x}x$ are in $\exp(\bigoplus_2^m V_j)$.*

Proof. By Campbell—Hausdorff, if $x=\exp(X+Y)$ as above, $x\tilde{x}=\exp(2Y+\text{commutators})\in\exp(\bigoplus_2^m V_j)$, and likewise for $\tilde{x}x$.

We now derive some important properties of the spaces Γ_α.

(5.4) **Proposition.** *There is a constant $C>0$ such that if $g\in\mathscr{BC}$ and $X_jg\in\mathscr{BC}$ for $j=1, \dots, n$, then*

$$\sup_{x,y} |g(xy)-g(x)|/|y| \le C\sum_1^n \|X_jg\|_\infty.$$

Proof. First suppose $y=\exp Y$ with $Y\in V_1$. Then $Yg\in\mathscr{BC}$, so

$$|g(xy)-g(x)| = \left|\int_0^1 Yg(x\exp(tY))\,dt\right| \le \|Yg\|_\infty.$$

Moreover, $ry=\exp(rY)$ for $r>0$, so

$$\|Yg\|_\infty \le |y|\sup\{\|Xg\|_\infty\colon X\in V_1,\ |\exp X|=1\} \le C|y|\sum_1^n \|X_jg\|_\infty.$$

Thus the assertion is true when y is restricted to $\exp V_1$. Next, given any $y \in G$, write $y = \prod_1^N y_j$ as in Lemma (5.1). Then

$$g(xy)-g(x) = [g(xy_1 \dots y_N)-g(xy_1 \dots y_{N-1})]+ \dots +[g(xy_1y_2)-g(xy_1)]+ \\ +[g(xy_1)-g(x)]$$

so that

$$|g(xy)-g(x)| \leqq C\left(\sum_1^N |y_j|\right)\left(\sum_1^n \|X_j g\|_\infty\right) \leqq NAC|y| \sum_1^n \|X_j g\|_\infty,$$

and the proof is complete.

In what follows it will sometimes be convenient to denote $f(xy)+f(xy^{-1})-2f(x)$ by $\Delta_y^2 f(x)$.

(5.5) **Proposition.** *Given* $0<\alpha<2$, *there is a constant* $C>0$ *such that if* $f \in \Gamma_\alpha$, *then*

(i) $\sup_{x,y} |f(xy)+f(xy^{-1})-2f(x)|/|y|^\alpha \leqq C|f|_\alpha$

(ii) $\sup_{x,y} |f(xy)+f(x\tilde{y})-2f(x)|/|y|^\alpha \leqq C|f|_\alpha$.

Proof. We first consider (i), which is trivial for $\alpha \leqq 1$. Suppose then that $f \in \Gamma_\alpha$, $1<\alpha<2$, and assume for the moment that f has compact support. By Lemma (4.12), $f = \sum_1^n g_j * K_j$ where $g_j = X_j f \in \Gamma_{\alpha-1}$ and K_j is a kernel of type 1. Then

$$\Delta_y^2 f(x) = \sum_1^n \int g_j(xz^{-1}) \Delta_y^2 K_j(z)\, dz. \tag{5.6}$$

By Proposition (1.15), there exists $\varepsilon>0$ such that $|\Delta_y^2 K_j(z)| \leqq C|y|^2 |z|^{-Q-1}$ if $|y| \leqq \varepsilon |z|$, and in particular $\Delta_y^2 K_j \in L^1$. We claim $\int \Delta_y^2 K_j = 0$. Indeed, if χ_r is the characteristic function of the set $\{z\colon |z| \leqq r\}$, we have $K_j \chi_r \in L^1$, so clearly $\int \Delta_y^2 (K_j \chi_r) = 0$. $\Delta_y^2(K_j \chi_r)$ converges pointwise and boundedly to $\Delta_y^2 K_j$ as $r \to \infty$, so by the Lebesgue convergence theorem $\int \Delta_y^2 K_j = 0$. Thus (5.6) can be rewritten as

$$\Delta_y^2 f(x) = \sum_1^n \int (g_j(xz^{-1})-g_j(x)) \Delta_y^2 K_j(z)\, dz. \tag{5.7}$$

Now by Corollary (1.6) and Proposition (1.15),

$$\left|\int_{\varepsilon|z| \geqq |y|} (g_j(xz^{-1})-g_j(x)) \Delta_y^2 K_j(z)\, dz\right| \leqq C|g_j|_{\alpha-1} \int_{\varepsilon|z| \geqq |y|} |z|^{\alpha-1} |y|^2 |z|^{-Q-1}\, dz \\ \leqq C'|g_j|_{\alpha-1} |y|^2 (|y|\varepsilon^{-1})^{\alpha-2} \leqq C'|g_j|_{\alpha-1} |y|^\alpha.$$

On the other hand,

$$\left|\int_{\varepsilon|z|<|y|} (g_j(xz^{-1})-g_j(x)) \Delta_y^2 K_j(z)\, dz\right| \leqq |g_j|_{\alpha-1} (\varepsilon^{-1}|y|)^{\alpha-1} \int_{\varepsilon|z|<|y|} |\Delta_y^2 K_j(z)|\, dz,$$

and since (by Proposition (1.4)) $\varepsilon|z| \leqq |y|$ implies $|zy| \leqq B|y|$ and $|zy^{-1}| \leqq B|y|$ for some $B \geqq \varepsilon^{-1}$,

$$\int_{\varepsilon|z| \leqq |y|} |\Delta_y^2 K_j(z)|\, dz \leqq 4 \int_{|z| \leqq B|y|} |K_j(z)|\, dz \leqq C \int_{|z| \leqq B|y|} |z|^{1-Q}\, dz \leqq C'(B|y|).$$

Combining these results with (5.7), we have

$$\sup_{x,y} |\Delta_y^2 f(x)|/|y|^\alpha \leqq C\sum_1^n |g_j|_{\alpha-1} \leqq C|f|_\alpha.$$

We now remove the restriction that f have compact support. If $f\in\Gamma_\alpha$ $(1<\alpha<2)$ and $\varphi\in\mathcal{D}$, it is easily verified that $|\varphi f|_\alpha \leqq C_0|\varphi|_\alpha|f|_\alpha$. Therefore, let $\{\varphi_j\}_1^\infty\subset\mathcal{D}$ be a partition of unity with the following properties: (a) $\sup_j |\varphi_j|_\alpha = C_1<\infty$, (b) if $U_j=\{x\colon x=zy$ with $z\in\operatorname{supp}\varphi_j$ and $|y|\leqq(4\|f\|_\infty/|f|_\alpha)^{1/\alpha}\}$, there is an integer N such that each $x\in G$ is contained in at most N U_j's. (We assume $f\neq$constant, so $|f|_\alpha\neq 0$.) Then $|\Delta_y^2 f(x)|\leqq\sum_1^\infty|\Delta_y^2(\varphi_j f)(x)|$. For each $x\in G$, the sum on the right contains at most N non-vanishing terms if $|y|\leqq(4\|f\|_\infty/|f|_\alpha)^{1/\alpha}$, so

$$\sup\{|\Delta_y^2 f(x)|/|y|^\alpha\colon x\in G,\ |y|\leqq(4\|f\|_\infty/|f|_\alpha)^{1/\alpha}\}\leqq NC_0C_1|f|_\alpha.$$

But $\quad |\Delta_y^2 f(x)|/|y|^\alpha \leqq 4\|f\|_\infty(|f|_\alpha/4\|f\|_\infty)\leqq|f|_\alpha\quad$ for $\quad |y|\geqq(4\|f\|_\infty/|f|_\alpha)^{1/\alpha}$.

so (i) is established.

Finally, we deduce (ii) from (i). If $Y\in V_j$ and $x\in G$, the function $f_{x,Y}\colon \mathbf{R}\to\mathbf{C}$ defined by $f_{x,Y}(t)=f(x\exp(tY))$ satisfies

$$|f_{x,Y}(t+s)+f_{x,Y}(t-s)-2f_{x,Y}(t)|\leqq C|f|_\alpha|\exp(sY)|^\alpha = C|f|_\alpha|\exp Y|^\alpha s^{\alpha/j}.$$

If $j\geqq 2$, then $\alpha/j<1$, so classical Lipschitz theory (cf. Stein [25]) implies that

$$|f_{x,Y}(t+s)-f_{x,Y}(t)|\leqq C|f|_\alpha|\exp Y|^\alpha s^{\alpha/j} = C|f|_\alpha|\exp(sY)|^\alpha.$$

Moreover, these estimates hold uniformly in x. Thus

$$|f(xy)-f(x)|\leqq C|f|_\alpha|y|^\alpha \quad (y\in\exp V_j,\ j\geqq 2).$$

By Lemma (5.2) and the collapsing-sum argument in the proof of Proposition (5.4), then,

$$|f(xy)-f(x)|\leqq C|f|_\alpha|y|^\alpha \quad (y\in\exp(\oplus_2^m V_j)).$$

In particular, by Lemma (5.3) and Proposition (1.4),

$$|f(xy^{-1})-f(x\tilde{y})|\leqq C|f|_\alpha|y\tilde{y}|^\alpha\leqq C'|f|_\alpha|y|^\alpha$$

since $|\tilde{y}|=|y|$. Thus

$$|f(xy)+f(x\tilde{y})-2f(x)|\leqq|\Delta_y^2 f(x)|+|f(xy^{-1})-f(x\tilde{y})|\leqq C''|f|_\alpha|y|^\alpha.$$

The proof is complete.

(5.8) **Proposition.** *Suppose* $0<\beta<1$ *and* $f\in\mathcal{BC}$. *Then* $f\in\Gamma_1$ *if and only if there exist* $C_0>0$ *and, for each* $\tau>0$, *functions* $f_\tau\in\Gamma_{1+\beta}$ *and* $f^\tau\in\Gamma_{1-\beta}$ *such that* $|f_\tau|_{1+\beta}\leqq C_0\tau$, $|f^\tau|_{1-\beta}\leqq C_0\tau^{-1}$, *and* $f=f_\tau+f^\tau$. *Moreover,* $C_0\sim|f|_1$.

Proof. Suppose that $f=f_\tau+f^\tau$ as above for all $\tau>0$. Then $\|\Delta_y^2 f^\tau\|_\infty \leqq 2C_0|y|^{1-\beta}\tau^{-1}$, while by Proposition (5.5), $\|\Delta_y^2 f_\tau\|_\infty \leqq CC_0|y|^{1+\beta}\tau$. Hence for all $\tau>0$,

$$\|\Delta_y^2 f\|_\infty \leqq C'C_0(|y|^{1-\beta}\tau^{-1}+|y|^{1+\beta}\tau).$$

Taking $\tau=|y|^{-\beta}$, we conclude that $\|\Delta_y^2 f\|_\infty \leqq C'C_0|y|$. Thus $f\in\Gamma_1$ and $|f|_1\leqq C'C_0$.

Conversely, suppose $f\in\Gamma_1$. Choose $\varphi_0\in\mathcal{D}$ supported in $\{x\colon |x|\leqq 1\}$ and satisfying $\varphi_0\geqq 0$, $\varphi_0(x)=\varphi_0(x^{-1})$, and $\int\varphi_0=1$. For $k\geqq 1$ set $\varphi_k(x)=2^{Qk}\varphi(2^k x)$, so that $\int\varphi_k=1$ for all k, and $\{\varphi_k\}$ is an approximation to the identity. Also, set $f_k=f*\varphi_k$ and $g_k=f_k-f_{k-1}$. Since f is continuous, we can write $f=f_0+\sum_1^\infty g_k$, the sum converging uniformly on compact sets. We claim that for some $C>0$ independent of k and f,

$$\|g_k\|_\infty \leqq C|f|_1 2^{-k}, \quad \|X_j g_k\|_\infty \leqq C|f|_1, \quad \|X_iX_jg_k\|_\infty \leqq C|f|_1 2^k \quad (i,j=1,\dots,n). \tag{5.9}$$

It suffices to estimate g_k and $X_iX_jg_k$, as an elementary argument then yields the estimate for X_jg_k. Since φ_k is even and $\int\varphi_k=\int\varphi_{k-1}$,

$$\begin{aligned} g_k(x) &= \int f(xy^{-1})\varphi_k(y)\,dy-\int f(xy^{-1})\varphi_{k-1}(y)\,dy \\ &= \frac{1}{2}\int [f(xy^{-1})+f(xy)-2f(x)][\varphi_k(y)-\varphi_{k-1}(y)]\,dy, \end{aligned}$$

whence

$$\|g_k\|_\infty \leqq \frac{1}{2}|f|_1\int_{|y|\leqq 2^{1-k}}|y||\varphi_k(y)-\varphi_{k-1}(y)|\,dy \leqq 2^{1-k}|f|_1\int\varphi_0=(2|f|_1)2^{-k}.$$

On the other hand, the function $\psi_k=X_iX_j(\varphi_k-\varphi_{k-1})$ satisfies $\psi_k(x)=\psi_k(\tilde{x})$ since the derivatives X_j reverse parity in the V_1 directions. Moreover, $\int\psi_k=0$ since ψ_k is the derivative of a function in $\mathcal{D}$. Thus

$$X_iX_jg_k(x)=\frac{1}{2}\int [f(xy^{-1})+f(x\tilde{y}^{-1})-2f(x)]\psi_k(y)\,dy.$$

Then by Proposition (5.5) and the fact that $\int|\psi_k|=2^{2k}\int|\psi_0|$,

$$\|X_iX_jg_k\|_\infty \leqq \frac{1}{2}|f|_1\int_{|y|\leqq 2^{1-k}}|y|\,|\psi_k(y)|\,dy \leqq 2^{-k}|f|_1\int|\psi_k| \leqq \left(\int|\psi_0|\right)|f|_1 2^k.$$

Thus (5.9) is established.

From (5.9) and Proposition (5.4), then,

$$\sup_x |g_k(xy)-g_k(x)| \leqq C|f|_1\min(2^{-k},|y|) \leqq C|f|_1 2^{-k\beta}|y|^{1-\beta},$$

$$\sup_x |X_jg_k(xy)-X_jg_k(x)| \leqq C|f|_1\min(1,2^k|y|) \leqq C|f|_1 2^{k\beta}|y|^{\beta}.$$

Thus $|g_k|_{1-\beta} \leqq C2^{-k\beta}|f|_1$ and $|g_k|_{1+\beta} \leqq C2^{k\beta}|f|_1$, so that

$$\left|f_0+\sum_1^N g_k\right|_{1+\beta} \leqq C|f|_1 \sum_1^N 2^{k\beta} \leqq C|f|_1 2^{N\beta},$$

$$\left|\sum_{N+1}^\infty g_k\right|_{1-\beta} \leqq C|f|_1 \sum_{N+1}^\infty 2^{-k\beta} \leqq C|f|_1 2^{-N\beta}.$$

Therefore, given $\tau \geqq 1$ we take $f_\tau = f_0+\sum_1^N g_k$ and $f^\tau = \sum_{N+1}^\infty g_k$ where $2^{(N-1)\beta} \leqq \tau < 2^{N\beta}$ and given $\tau<1$ we take $f_\tau=0$ and $f^\tau=f=f_0+\sum_1^\infty g_k$, and we are done.

(5.10) **Proposition.** *If $\alpha>\beta$ then $\Gamma_\alpha \subset \Gamma_\beta$.*

Proof. It is clear that $\Gamma_\alpha \subset \Gamma_\beta$ when the interval $[\alpha, \beta]$ does not contain an integer. Proposition (5.4) shows that $\Gamma_\alpha \subset \Gamma_\beta$ when $\beta \leqq 1$ and $1<\alpha<2$, and this combined with Proposition (5.8) shows that $\Gamma_1 \subset \Gamma_\beta$ for $0<\beta<1$. The assertion is thus proved for $0<\beta<\alpha<2$, and the general case follows by applying these arguments to derivatives.

(5.11) **Proposition.** *If $1 \leqq p < \infty$ and $0<\alpha\leqq 1$, there is a constant $C=C_{p,\alpha}>0$ such that $\|f\|_\infty \leqq C(\|f\|_p+|f|_\alpha)$ for all $f \in L^p \cap \Gamma_\alpha$.*

Proof. We may assume $\alpha<1$, since by Proposition (5.10) (and its proof), $\Gamma_1 \subset \Gamma_\alpha$ and $|f|_\alpha \leqq C|f|_1$ for $\alpha<1$. For any $x \in G$ we have $|f(y)| \geqq \frac{1}{2}|f(x)|$ for all y such that $|xy^{-1}| \leqq (|f(x)|/2|f|_\alpha)^{1/\alpha} = A$. Thus

$$\|f\|_p^p \geqq \int_{|xy^{-1}|\leqq A} |f(y)|^p\,dy \geqq \left(\frac{1}{2}|f(x)|\right)^p \cdot CA^Q = C'|f(x)|^{p+(Q/\alpha)}|f|_\alpha^{-Q/\alpha},$$

or, with $\gamma = Q/\alpha p$,

$$\|f\|_\infty = \sup_x |f(x)| \leqq C\|f\|_p^{1/(1+\gamma)}|f|_\alpha^{\gamma/(1+\gamma)} \leqq C(\|f\|_p+|f|_\alpha).$$

We now come to the main topic of this section: the effect of convolution with kernels of type λ $(0\leqq\lambda<Q)$ on the spaces Γ_α. Actually, if $k<\alpha\leqq k+1$ where k is an integer, we shall consider not Γ_α itself but $S_k^p \cap \Gamma_\alpha$ $(1<p<\infty)$, the space of functions f such that f and $X_I f$ are in $L^p \cap \Gamma_{\alpha-k}$ for $|I|\leqq k$, in order to guarantee that the integrals in question converge. We note that by Proposition (5.11), $S_k^p \cap \Gamma_\alpha$ is a Banach space with norm $\|f\|_{p,k}+|f|_\alpha$.

(5.12) **Theorem.** *Let K be a kernel of type 0, $k=0, 1, 2, \ldots$, $k<\alpha\leqq k+1$, and $1<p<\infty$. Then the mapping $T: f \to f*K$ is bounded on $S_k^p \cap \Gamma_\alpha$.*

Proof. The case $0<\alpha<1$ is due to Korányi—Vági [21], and their argument shows that $|Tf|_\alpha \leqq C|f|_\alpha$ for $f\in\Gamma_\alpha$, $0<\alpha<1$. We refer to their paper for the demonstration.

Suppose $1<\alpha<2$. If $f \in S_1^p \cap \Gamma_\alpha$, we know by Theorem (4.9) and the result for $\alpha<1$ that $Tf \in S_1^p \cap \Gamma_{\alpha-1}$ and

$$\|Tf\|_{p,1}+|Tf|_{\alpha-1} \leqq C(\|f\|_{p,1}+|f|_{\alpha-1}).$$

On the other hand, if $f\in\mathcal{D}$ we can write $f=\sum_1^n X_i f*K_i$ where K_i is a kernel of type 1, by Lemma (4.12). Then by Proposition (1.13),

$$X_j Tf = X_j \sum_{i=1}^n X_i f*(K_i*K) = \sum_{i=1}^n X_i f*K_{ij}$$

where $K_{ij}=X_j(K_i*K)$ is a kernel of type 0. Since $\mathcal{D}$ is dense in S_1^p and $g\to g*K_{ij}$ is bounded on L^p, this equation remains valid for all $f\in S_1^p$. In particular, if $f\in S_1^p\cap\Gamma_\alpha$, $1<\alpha<2$, then $X_i f\in L^p\cap\Gamma_{\alpha-1}$; therefore $X_j Tf\in L^p\cap\Gamma_{\alpha-1}$ and

$$|X_j Tf|_{\alpha-1} \leqq C\sum_1^n |X_i f|_{\alpha-1} \leqq C|f|_\alpha.$$

The theorem is thus established for $0<\alpha<1$ and $1<\alpha<2$. Next, if $f\in\Gamma_1\cap L^p$, by Proposition (5.8) we can write $f=f_\tau+f^\tau$ with $|f_\tau|_{3/2}\leqq C|f|_1\tau$ and $|f^\tau|_{1/2}\leqq C|f|_1\tau^{-1}$, for every $\tau>0$. Moreover, the proof of Proposition (5.8) shows that we can take $f_\tau=f*\varphi$ for some $\varphi\in\mathcal{D}$, so that $f_\tau\in S_1^p$ and $f^\tau\in L^p$. By the preceding results, then, $|Tf_\tau|_{3/2}\leqq C'|f|_1\tau$ and $|Tf^\tau|_{1/2}\leqq C'|f|_1\tau^{-1}$, so by the converse part of Proposition (5.8) $Tf\in\Gamma_1$ and $|Tf|_1\leqq C''|f|_1$.

The theorem is therefore true for $0<\alpha<2$, and the general case now follows easily by induction on k by using the kernels K_{ij} as in the proof for $1<\alpha<2$.

As a consequence, we deduce the following boundedness theorem for kernels of higher type.

(5.13) **Theorem.** *Let K be a kernel of type λ, $\lambda=1, 2, \dots, Q-1$, and suppose $1<p<q<\infty$ and $q^{-1}=p^{-1}-(\lambda/Q)$. If $k=0, 1, 2, \dots$ and $k<\alpha\leqq k+1$, the mapping $T\colon f\to f*K$ is bounded from $S_k^p\cap\Gamma_\alpha$ to $S_{k+\lambda}^q\cap\Gamma_{\alpha+\lambda}$.*

Proof. First suppose $\lambda=1$. If $f\in S_k^p\cap\Gamma_\alpha$ with p, q, k, α as above, then $Tf\in L^q$ and $\|Tf\|_q\leqq C\|f\|_p$ by Proposition (1.11). Also, if we set $K_0(x)=K(x)$ when $|x|\leqq 1$ and $=0$ otherwise, and $K_\infty=K-K_0$, we have $K_0\in L^1$ and $K_\infty\in L^{p'}$ where $p'=p/(p-1)$, so by Proposition (5.11),

$$\|Tf\|_\infty \leqq \|f\|_\infty\|K_0\|_1+\|f\|_p\|K_\infty\|_{p'} \leqq C(\|f\|_p+|f|_\alpha).$$

This shows that Tf is bounded, and also that it is continuous since $\mathcal{D}$ is dense in $S_k^p\cap\Gamma_\alpha$. Moreover, $X_j Tf=f*X_j K$, and since $X_j K$ is a kernel of type 0, by Theorem (5.12) we have $X_j Tf\in S_k^p\cap\Gamma_\alpha\subset S_k^q\cap\Gamma_\alpha$ since $L^p\cap L^\infty\subset L^q$, and

$$\|X_j Tf\|_{q,k}+|X_j Tf|_\alpha \leqq C(\|X_j Tf\|_{p,k}+|X_j Tf|_\alpha) \leqq C'(\|f\|_{p,k}+|f|_\alpha).$$

By Proposition (5.4), $Tf\in\Gamma_\beta$ for $0<\beta\leqq 1$; therefore $Tf\in S_{k+1}^q\cap\Gamma_{\alpha+1}$ and

$$\begin{aligned}\|Tf\|_{q,k+1}+|Tf|_{\alpha+1} &\leqq \|Tf\|_q+|Tf|_{\alpha-k}+\sum_1^n(\|X_j Tf\|_{q,k}+|X_j Tf|_\alpha)\\ &\leqq C(\|f\|_{p,k}+|f|_\alpha).\end{aligned}$$

The theorem now follows by induction on λ. If K is a kernel of type λ ($\lambda=2, 3, \dots, Q-1$) and $f\in S_k^p\cap\Gamma_\alpha$, we have $X_j Tf=f*X_j K$ where $X_j K$ is a kernel of type $\lambda-1$. By inductive hypothesis, $X_jTf\in S^r_{k+\lambda-1}\cap\Gamma_{\alpha+\lambda-1}\subset S^q_{k+\lambda-1}\cap\Gamma_{\alpha+\lambda-1}$ where $r^{-1}=p^{-1}-(\lambda-1)/Q$, and by the preceding argument $Tf\in L^q\cap\Gamma_\beta$ for $0<\beta\leqq 1$, hence $Tf\in S^q_{k+\lambda}\cap\Gamma_{\alpha+\lambda}$. (The norm inequalities are easily checked.)

(5.14) **Theorem.** *Suppose λ, p, r are numbers satisfying $0<\lambda<Q$, $1<p<(Q/\lambda)<r\leqq\infty$. Define q, α by $q^{-1}=p^{-1}-(\lambda/Q)$ and $\alpha=\lambda-(Q/r)$, so $p<q<\infty$ and $0<\alpha\leqq\lambda$. If K is a kernel of type λ, the mapping $T\colon f\to f*K$ is bounded from $L^p\cap L^r$ to $L^q\cap\Gamma_\alpha$. Moreover, for $0<\alpha\leqq 1$ there is a constant $C>0$ such that $|Tf|_\alpha\leqq C\|f\|_r$.*

Proof. We know by Proposition (1.11) that T is bounded from L^p to L^q. To prove the theorem for the case $0<\alpha\leqq 1$, then, it suffices to show that $|Tf|_\alpha\leqq C\|f\|_r$.

First suppose $0<\alpha<1$. We have

$$Tf(xy)-Tf(x)=\int f(xz^{-1})[K(zy)-K(z)]\,dz.$$

We shall estimate the integral over the regions $|z|>2|y|$ and $|z|\leqq 2|y|$ separately. For the first one, we note that if $r'=r/(r-1)$, then $\lambda-Q+(Q/r')=\lambda-(Q/r)=\alpha$. In particular, $(\lambda-Q-1)r'=(\alpha-1)r'-Q<-Q$, so by Proposition (1.15) and Hölder's inequality,

$$\Big|\int_{|z|>2|y|}f(xz^{-1})[K(zy)-K(z)]\,dz\Big|\leqq C\|f\|_r\Big(\int_{|z|>2|y|}[|y|\,|z|^{\lambda-Q-1}]^{r'}\,dz\Big)^{1/r'}$$
$$\leqq C'\|f\|_r|y|\,(2|y|)^{\lambda-Q-1+(Q/r')}\leqq C'\|f\|_r|y|^\alpha.$$

On the other hand, by Proposition (1.4) there is a constant $B\geqq 2$ such that $|z|\leqq 2|y|$ implies $|zy|\leqq B|y|$, so by Hölder again,

$$\Big|\int_{|z|\leqq 2|y|}f(xz^{-1})[K(zy)-K(z)]\,dz\Big|\leqq\|f\|_r\Big(2\int_{|z|\leqq B|y|}|K(z)|^{r'}\,dz\Big)^{1/r'}$$
$$\leqq C\|f\|_r\Big(\int_{|z|\leqq B|y|}|z|^{(\lambda-Q)r'}\,dz\Big)^{1/r'}\leqq C'\|f\|_r(B|y|)^{\lambda-Q+(Q/r')}\leqq C''\|f\|_r|y|^\alpha.$$

Therefore $|Tf|_\alpha\leqq C\|f\|_r$ for $0<\alpha<1$.

Next suppose $\alpha=1$. Here we have

$$Tf(xy)+Tf(xy^{-1})-2Tf(x)=\int f(xz^{-1})[K(zy)+K(zy^{-1})-2K(z)]\,dz.$$

The estimate $|Tf|_1\leqq C\|f\|_r$ then follows by the same argument as above, using the estimate for second differences in Proposition (1.15).

For the general case, write $\alpha=k+\alpha'$ where $k=0,1,2,\ldots$ and $0<\alpha'\leqq 1$. The theorem is proved for $k=0$; we assume $k\geqq 1$ and proceed by induction. Noting that $X_jTf=f*X_jK$ and X_jK is a kernel of type $\lambda-1>0$, we have $X_jTf\in L^s\cap\Gamma_{\alpha-1}$ where

$s^{-1}=p^{-1}-(\lambda-1)/Q$ by inductive hypothesis. Hence

$$X_j Tf \in (L^s \cap L^\infty) \cap \Gamma_{\alpha-1} \subset L^q \cap \Gamma_{\alpha-1}, \quad \text{and} \quad \|X_j Tf\|_{\Gamma_{\alpha-1}} \leqq C(\|f\|_p + \|f\|_r).$$

But, setting $K=K_0+K_\infty$ as in the proof of Theorem (5.13), we conclude by the argument given there that $Tf \in \mathscr{BC}$ and

$$\|Tf\|_\infty \leqq \|f\|_p \|K_0\|_{p'} + \|f\|_r \|K_\infty\|_{r'} \leqq C(\|f\|_p + \|f\|_r).$$

Hence, in view of Proposition (5.4),

$$\|Tf\|_{\Gamma_\alpha} \leqq \|Tf\|_{\Gamma_{\alpha'}} + \textstyle\sum_1^n \|X_j Tf\|_{\Gamma_{\alpha-1}} \leqq \|Tf\|_\infty + C \sum_1^n \|X_j Tf\|_{\Gamma_{\alpha-1}}$$

$$\leqq C'(\|f\|_p + \|f\|_r),$$

and we are done.

The following theorem generalizes the famous Sobolev imbedding theorem (the classical version being the case $G=\mathbf{R}^n$).

(5.15) **Theorem.** *Suppose* $1<p<\infty$ *and* $\alpha>Q/p$. *Then* $S_\alpha^p \subset \Gamma_\beta$ *and* $\| \ \|_{\Gamma_\beta} \leqq C \| \ \|_{p,\alpha}$ *where* $\beta=\alpha-(Q/p)$.

Proof. First consider the case where $p>Q/(Q-1)$ and $\beta \leqq 1$, which implies $\alpha<Q$. Since $\mathscr{D}$ is dense in S_α^p (Theorem (4.5)), it suffices to prove the estimate $\|u\|_{\Gamma_\beta} \leqq C\|u\|_{p,\alpha}$ for $u \in \mathscr{D}$. Now $\mathscr{D} \subset S_\alpha^q$ for all q, so for $u \in \mathscr{D}$, $\mathscr{J}^{\alpha/2}u \in L^q$ for all q. By Propositions (1.11) and (3.18), then, $u=\mathscr{J}^{-\alpha/2}\mathscr{J}^{\alpha/2}u=(\mathscr{J}^{\alpha/2}u) * R_\alpha$. Therefore, by Theorem (5.14) (with r replaced by p and p replaced by some $q<Q/\alpha$), we have $|u|_\beta \leqq C\|\mathscr{J}^{\alpha/2}u\|_p$. Then by Proposition (5.11),

$$\|u\|_{\Gamma_\beta} \leqq C'(|u|_\beta + \|u\|_p) \leqq C''(\|\mathscr{J}^{\alpha/2}u\|_p + \|u\|_p) = C''\|u\|_{p,\alpha}.$$

Next, suppose $\beta \leqq 1$ and $p \leqq Q/(Q-1)$. Choose a number q with $Q/(Q-1)<q<\infty$, and set $\gamma=\alpha-Q(p^{-1}-q^{-1})$. Clearly $\gamma=\beta+(Q/q)>0$. By the preceding remarks together with Theorem (4.17), therefore, $S_\alpha^p \subset S_\gamma^q \subset \Gamma_\beta$ and $\| \ \|_{\Gamma_\beta} \leqq C\| \ \|_{q,\gamma} \leqq C'\| \ \|_{p,\alpha}$.

Finally, suppose $\beta>1$; we write $\beta=k+\beta'$ where $k=1, 2, 3, \ldots$ and $0<\beta' \leqq 1$. By Theorem (4.10), if $f \in S_\alpha^p$ then f and $X_I f$ are in $S_{\alpha-k}^p$ for $|I| \leqq k$, hence in $\Gamma_{\beta'}$. Thus $f \in \Gamma_\beta$, and the norm estimate is obvious.

(5.16) **Corollary.** $S_\alpha^p \subset \mathscr{C}^k$ *provided* $\alpha>mk+(Q/p)$, *where* m *is the number of steps in the stratification of* G.

We note that this result is sharper than the one obtained by combining Theorem (4.16) with the ordinary Sobolev theorem, namely $S_\alpha^p \subset L_{\alpha/m}^p(\text{loc}) \subset \mathscr{C}^k$ provided $\alpha>m(k+(N/p))$, where N is the Euclidean dimension of G.

Our last objective in this chapter is to compare the spaces Γ_α with the classical spaces Λ_α. To begin with, suppose $f\in\Gamma_\alpha$ has compact support and $0<\alpha<1$. By Lemma (1.3),

$$|f(xz)-f(x)| = O(|z|^\alpha) = O(\|z\|^{\alpha/m}) \quad \text{as} \quad z\to 0,$$

where m is the number of steps in the stratification. Thus if we set $z=x^{-1}(x+y)$,

$$|f(x+y)-f(x)| = |f(xz)-f(x)| = O(\|z\|^{\alpha/m}) = O(\|y\|^{\alpha/m}) \quad \text{as} \quad y\to 0.$$

Moreover, these estimates are uniform in x as x ranges over supp f, so we conclude that $f\in\Lambda_{\alpha/m}$. Hence $\Gamma_\alpha(\text{loc})\subset\Lambda_{\alpha/m}(\text{loc})$ for $0<\alpha<1$ (where, as usual, $\Gamma_\alpha(\text{loc})=\{f\colon \varphi f\in\Gamma_\alpha \text{ for all } \varphi\in\mathcal{D}\}$ and likewise for $\Lambda_{\alpha/m}(\text{loc})$). Similarly, we see that $\Lambda_\alpha(\text{loc})\subset\Gamma_\alpha(\text{loc})$ for $0<\alpha<1$. These inclusions are best possible, as $|y|=\|y\|$ for $y\in\exp V_1$ and $|y|=\|y\|^{1/m}$ for $y\in\exp V_m$ by (1.17).

We shall show that in fact $\Lambda_\alpha(\text{loc})\subset\Gamma_\alpha(\text{loc})\subset\Lambda_{\alpha/m}(\text{loc})$ for all $\alpha>0$ and thus provide a result for Γ_α parallel to Theorem (4.16) for S^p_α. We assume throughout that $m>1$, as the Abelian case is trivial. The following line of proof was suggested to us by E. M. Stein.

(5.17) **Lemma.** *Suppose* $0<\alpha<2$ *and that* $f\in\mathcal{BC}$ *satisfies*

$$f(xy)+f(xy^{-1})-2f(x) = O(\|y\|^\alpha) \quad \text{as} \quad y\to 0$$

for x in any bounded set. Then $f\in\Lambda_\alpha(\text{loc})$.

Proof. By a well-known characterization of Λ_α, $0<\alpha<2$ (cf. Stein [25]), it suffices to show that

$$f(x+y)+f(x-y)-2f(x) = O(\|y\|^\alpha) \quad \text{as} \quad y\to 0.$$

If $Y\in\mathfrak{g}$, the hypothesis implies that the function $f_{x,Y}\colon\mathbf{R}\to\mathbf{C}$ defined by $f_{x,Y}(f)=f(x\exp(tY))$ is in $\Lambda_\alpha(\mathbf{R},\text{loc})$ for each x, hence in $\Lambda_{\alpha/2}(\mathbf{R},\text{loc})$, and the Lipschitz constants involved vary continuously with x and Y. Since $\alpha/2<1$, by taking $Y=\exp^{-1}(x^{-1}y)$ we see that

$$|f(y)-f(x)| = O(\|x^{-1}y\|^{\alpha/2}) = O(\|x-y\|^{\alpha/2}) \tag{5.18}$$

as $x-y\to 0$ and x, y range over a bounded set.

Now, for fixed $x\in G$, define $\psi(y)=x^{-1}(x+y)$. ψ is the inverse of the diffeomorphism φ of Lemma (1.14), which says that $y\to\varphi(y^{-1})$ has the same differential at 0 as $y\to\varphi(y)^{-1}$. Hence $y\to\psi(y^{-1})$ has the same differential at 0 as $y\to\psi(y)^{-1}$, and so

$$\|\psi(y)\| = O(\|y\|) \quad \text{and} \quad \|\psi(y^{-1})-\psi(y)^{-1}\| = O(\|y\|^2).$$

Thus, in view of (5.18),

$$\begin{aligned} f(x+y)+f(x-y)-2f(x) &= [f(x\psi(y))+f(x\psi(y)^{-1})-2f(x)] \\ &\quad +[f(x\psi(y^{-1}))-f(x\psi(y)^{-1})] \\ &= O(\|\psi(y)\|^{\alpha})+O(\|\psi(y^{-1})-\psi(y)^{-1}\|^{\alpha/2}) \\ &= O(\|y\|^{\alpha}). \end{aligned}$$

(5.19) **Lemma.** *If $0<\alpha<2$ and $f\in\Gamma_\alpha$ has compact support, then $f\in\Lambda_{\alpha/m}$.*

Proof. This is clear from Proposition (5.5) and Lemmas (1.3) and (5.17).

(5.20) **Lemma.** *Let K be a kernel of type 1. There is a constant $C>0$ such that for all $x\in G$ and $y\in\exp V_j$ $(j=1, \ldots, m)$ with $|y|\leqq\frac{1}{2}|x|$, $|K(xy)-K(x)|\leqq \leqq C\|y\|\,|x|^{1-Q-j}$.*

Proof. Since $\|ry\|=r^j\|y\|$ for $y\in\exp V_j$, both sides of the inequality are multiplied by r^{1-Q} when x and y are replaced by rx and ry. It therefore suffices to assume $|x|=1$ and $|y|\leqq\frac{1}{2}$, in which case the assertion is evident from the mean value theorem.

(5.21) **Lemma.** *If K is a kernel of type 1 and U is any bounded subset of G,*

$$\int_U |K(xy)-K(x)|\,dx = O(\|y\|^{1/m}) \quad \text{as} \quad y\to 0.$$

Proof. By Lemma (5.2), any $y\in G$ can be written $y=\prod_1^m y_j$ with $y_j\in\exp V_j$, and $\|y_j\|=O(\|y\|)$ as $y\to 0$ since the mapping $(y_1, \ldots, y_m)\to\prod_1^m y_j$ is a diffeomorphism. We then have

$$\begin{aligned} K(xy)-K(x) &= [K(xy_1\ldots y_m)-K(xy_1\ldots y_{m-1})]+\ldots+ \\ &\quad +[K(xy_1y_2)-K(xy_1)]+[K(xy_1)-K(x)]. \end{aligned}$$

If $x\in U$ and y is small, the points $xy_1\ldots y_j$ range over a bounded set, so it suffices to prove the estimate for $y\in\exp V_j$, $j=1, \ldots, m$.

We write

$$\int_U |K(xy)-K(x)|\,dx = I_1+I_2$$

where I_1 is the integral over $U\cap\{x\colon |x|\leqq 2|y|\}$ and I_2 is the integral over $U\cap\{x\colon |x|>2|y|\}$. By Proposition (1.4), $|x|\leqq 2|y|$ implies $|xy|\leqq B|y|$ for some $B\geqq 2$, so

$$I_1 \leqq 2\int_{|x|\leqq B|y|}|K(x)|\,dx \leqq C\int_{|x|\leqq B|y|}|x|^{1-Q}\,dx = O(|y|).$$

However, by (1.17) $|y|=\|y\|^{1/j}$ for $y\in\exp V_j$, so $I_1=O(\|y\|^{1/j})=O(\|y\|^{1/m})$.

To estimate I_2 we use Lemma (5.20):

$$I_2 \leqq C\|y\| \int_{U\cap\{x:|x|>2|y|\}} |x|^{1-Q-j}\,dx,$$

so that

$$I_2 = O(\|y\||y|^{1-j}) = O(\|y\|^{1/j}) \quad \text{if} \quad j>1,$$

$$I_2 = O\big(\|y\|\log(|y|^{-1})\big) = O\big(\|y\|\log(\|y\|^{-1})\big) \quad \text{if} \quad j=1.$$

In any event, $I_2 = O(\|y\|^{1/m})$ (since we are assuming $m>1$), so we are done.

(5.22) **Lemma.** *Suppose* $0<\beta<1$ *and* $f\in\Lambda_\beta$ *has compact support. If* K *is a kernel of type* 1 *then* $F=f*K$ *is in* $\Lambda_{\beta+(1/m)}(\text{loc})$.

Proof. Since $\beta+(1/m)<2$, by Lemma (5.17) it suffices to show that

$$\Delta_y^2 F(x) = F(xy)+F(xy^{-1})-2F(x) = O(\|y\|^{\beta+(1/m)}) \quad \text{as} \quad y\to 0.$$

In fact, we have

$$\Delta_y^2 F(x) = \int [f(xyz^{-1})-f(xz^{-1})][K(z)-K(zy^{-1})]\,dz.$$

For $\|y\|\leqq 1$ and $\|x\|$ bounded, the set $U=\{z: f(xyz^{-1})-f(xz^{-1})\neq 0\}$ is bounded, so by Lemma (5.21),

$$|\Delta_y^2 F(x)| \leqq \sup_{z\in U} |f(xyz^{-1})-f(xz^{-1})| \int_U |K(z)-K(zy^{-1})|\,dz$$

$$\leqq C\sup_{z\in U} \|zyz^{-1}\|^{\beta}\|y\|^{1/m} \leqq C'\|y\|^{\beta+(1/m)}.$$

(5.23) **Lemma.** *If* $0<\alpha\leqq m$ *and* $f\in\Gamma_\alpha$ *has compact support, then* $f\in\Lambda_{\alpha/m}$.

Proof. By Lemma (5.19) we may assume that $\alpha>1$, hence that $\alpha=k+\alpha'$ where $k=1,2,\dots,m-1$, and $0<\alpha'\leqq 1$. By Lemma (4.12), if $f\in\Gamma_\alpha$ has compact support, we have

$$f = \textstyle\sum_i (X_i f)*K_i$$

$$X_j f = \textstyle\sum_i (X_i X_j f)*K_i \tag{5.24}$$

$$\vdots$$

$$X_{j_1}\dots X_{j_{k-1}} f = \textstyle\sum_i (X_i X_{j_1}\dots X_{j_{k-1}} f)*K_i$$

where $K_1,\dots,K_n$ are kernels of type 1. Now $X_i X_{j_1}\dots X_{j_{k-1}} f$ is in $\Gamma_{\alpha'}$, hence in $\Lambda_{\alpha'/m}$ by Lemma (5.19). Therefore by Lemma (5.22), $X_{j_1}\dots X_{j_{k-1}} f\in\Lambda_{(\alpha'+1)/m}$ (globally, since it has compact support). Applying Lemma (5.22) repeatedly to the sequence of equations (5.24), we conclude that $X_{j_1}\dots X_{j_l} f\in\Lambda_{(\alpha'+l)/m}$, and finally — since $(\alpha'+k-1)/m<1$ — that $f\in\Lambda_{(\alpha'+k)/m}=\Lambda_{\alpha/m}$.

(5.25) **Theorem.** $\Lambda_\alpha(\text{loc})\subset\Gamma_\alpha(\text{loc})\subset\Lambda_{\alpha/m}(\text{loc})$ *for all* $\alpha>0$.

Proof. It suffices to show that if $f\in\Lambda_\alpha$ has compact support then $f\in\Gamma_\alpha$, and if $f\in\Gamma_\alpha$ has compact support then $f\in\Lambda_{\alpha/m}$. The first assertion is obvious in view of the fact that $\|x\|=O(|x|)$ as $x\to 0$. For the second, write $\alpha=mk+\alpha'$ where $k=0, 1, 2, \dots$ and $0<\alpha'\leqq m$. By Lemma (5.23), the theorem is proved for $k=0$. If $k\geqq 1$, then f and $X_I f$ are in $\Gamma_{\alpha'}$ for $|I|\leqq mk$, so that *all* k-th order derivatives of f are in $\Gamma_{\alpha'}$, hence in $\Lambda_{\alpha'/m}$. Therefore $f\in\Lambda_{k+(\alpha'/m)}=\Lambda_{\alpha/m}$, and the proof is complete.

It would be of interest to generalize some of the other aspects of the theory of differentiability on Euclidean spaces (e.g., Besov spaces, Poisson integrals, relationship between Bessel potentials and Lipschitz conditions: see Stein [25]) to the setting of stratified groups. However, these questions are beyond the scope of the present paper.

6. Regularity of homogeneous hypoelliptic operators

We conclude by applying the preceding material to obtain sharp L^p and Lipschitz regularity properties for homogeneous hypoelliptic operators on stratified groups. This theorem should be a prototype for regularity results for a much wider class of differential operators: see, for example, Folland—Stein [6], [7] where some special cases of this theorem are extended to non-homogeneous situations.

(6.1) **Theorem.** *Let G be a stratified group of homogeneous dimension $Q>2$, and let $\mathscr{L}$ be a left-invariant homogeneous differential operator of degree k, $0<k<Q$ (k is necessarily an integer) which is hypoelliptic together with its transpose $\mathscr{L}^t$. Suppose $U\subset G$ is an open set, and suppose $f, g\in\mathscr{D}'(U)$ satisfy $\mathscr{L}f=g$ on U. If $g\in S^p_\alpha(U,\text{loc})$ $(1<p<\infty,\ \alpha\geqq 0)$ then $f\in S^p_{\alpha+k}(U,\text{loc})$, and if $g\in\Gamma_\alpha(U,\text{loc})$ then $f\in\Gamma_{\alpha+k}(U,\text{loc})$.*

Proof. Let K_0 be the fundamental solution for $\mathscr{L}$ given by Theorem (2.1). Given any compact set $W\subset U$, choose $\varphi\in\mathscr{D}(U)$ with $\varphi=1$ on a neighborhood of W and set $u=(\varphi g)*K_0$. Then $\mathscr{L}(f-u)=(1-\varphi)g$, so by hypoellipticity of $\mathscr{L}$, $f-u$ is $\mathscr{C}^\infty$ on a neighborhood of W. It therefore suffices to show that u has the required regularity properties.

First, if $g\in\Gamma_\alpha(U,\text{loc})$, it follows from Theorem (5.13) that $u\in\Gamma_{\alpha+k}$. On the other hand, suppose $g\in S^p_\alpha(U,\text{loc})$. We have $X_I u=(\varphi g)*X_I K_0$, and $X_I K_0$ is a kernel of type $k-|I|$. Since $\varphi g\in L^1\cap S^p_\alpha$, we see from Proposition (1.11) and Theorem (4.9) that u and $X_I u(|I|\leqq k)$ are in various L^q spaces with $q\geqq p$, hence in $L^p(\text{loc})$, and moreover that $X_I u\in S^p_\alpha$ for $|I|\neq k$. In view of Theorem (4.10), we have $u\in S^p_k(\text{loc})$ and $X_I u\in S^p_\alpha$ for $|I|=k$, and thus $u\in S^p_{\alpha+k}(\text{loc})$. The proof is complete.

One can obtain other regularity results for $\mathscr{L}$ in terms of S^p_α and Γ_α or the classical spaces L^p_α, Λ_α, and $\mathscr{C}^k$ by combining this theorem with Theorems (4.16), (4.17), (5.15), (5.25), and Corollary (5.16).

References

1. Bony, J. M., Principe du maximum, inégalité de Harnack, et unicité du probléme de Cauchy pour les operateurs elliptiques dégénérés, *Ann. Inst. Fourier Grenoble,* **19** (1) (1969), 277—304.
2. Calderón, A. P., Lebesgue spaces of differentiable functions and distributions, *Proc. Symp. Pure Math.,* **4** (1961), 33—49.
3. Coifman, R. and Weiss, G. *Analyse harmonique non-commutative sur certains espaces homogènes,* Lecture notes #242, Springer-Verlag, Berlin, (1971).
4. Dyer, J. L. A nilpotent Lie algebra with nilpotent automorphism group, *Bull. Amer. Math. Soc.,* **76** (1970), 52—56.
5. Folland, G. B. and Kohn, J. J. *The Neumann problem for the Cauchy—Riemann complex,* Ann. of Math. Studies #75, Princeton University Press, Princeton, (1972).
6. Folland, G. B. and Stein, E. M. Parametrices and estimates for the $\bar{\partial}_b$ complex on strongly pseudoconvex boundaries, *Bull. Amer. Math. Soc.,* **80** (1974), 253—258.
7. Folland, G. B. and Stein, E. M. Estimates for the $\bar{\partial}_b$ complex and analysis on the Heisenberg group, *Comm. Pure Appl. Math.,* **27** (1974), 429—522.
8. Guillemin, V. and Sternberg, S. Subelliptic estimates for complexes, *Proc. Nat. Acad. Sci. U.S.A.,* **67** (1970), 271—274.
9. Hochschild, G. *The structure of Lie groups,* Holden-Day, San Francisco, (1965).
10. Hörmander, L. Hypoelliptic second-order differential equations, *Acta Math.,* **119** (1967), 147—171.
11. Hunt, G. A. Semigroups of measures on Lie groups, *Trans. Amer. Math. Soc.,* **81** (1956), 264—293.
12. Jørgensen, P. Representations of differential operators on a Lie group, to appear.
13. Knapp, A. W. and Stein, E. M. Intertwining operators for semi-simple groups, *Ann. of Math.,* **93** (1971), 489—578.
14. Kohn, J. J. and Nirenberg, L. Non-coercive boundary value problems, *Comm. Pure Appl. Math.,* **18** (1965), 443—492.
15. Komatsu, H. Fractional powers of operators, *Pac. J. Math.,* **19** (1966), 285—346.
16. Komatsu, H. Fractional powers of operators, II: Interpolation spaces, *Pac. J. Math.,* **21** (1967), 89—111.
17. Komatsu, H. Fractional powers of operators, III: Negative powers, *J. Math. Soc. Japan,* **21** (1969), 205—220.
18. Komatsu, H. Fractional powers of operators, IV: Potential operators, *J. Math. Soc. Japan,* **21** (1969), 221—228.
19. Komatsu, H. Fractional powers of operators, V: Dual operators, *J. Fac. Sci. Univ. Tokyo,* Sec. IA, **17** (1970), 373—396.
20. Komatsu, H. Fractional powers of operators, VI: Interpolation of nonnegative operators and imbedding theorems, *J. Fac. Sci. Univ. Tokyo,* Sec. IA, **19** (1972), 1—62.
21. Korányi, A. and Vági, S. Singular integrals in homogeneous spaces and some problems of classical analysis, *Ann. Scuola Norm. Sup. Pisa,* **25** (1971), 575—648.
22. Oleĭnik, O. A. and Radkevič, E. V. *Second order equations with nonnegative characteristic form,* Amer. Math. Soc., Providence, (1973).
23. Schwartz, L. *Théorie des distributions,* Hermann, Paris, (1966).
24. Stein, E. M. *Topics in harmonic analysis,* Ann. of Math. Studies #63, Princeton University Press, Princeton, (1970).

25. STEIN, E. M. *Singular integrals and differentiability properties of functions,* Princeton University Press, Princeton, (1970).
26. STEIN, E. M. Some problems in harmonic analysis suggested by symmetric spaces and semi-simple groups, *Proc. Internat. Congress Math. Nice* (1970), vol. I, 173—189.
27. STEIN, E. M. Singular integrals and estimates for the Cauchy—Riemann equations, *Bull. Amer. Math. Soc.,* **79** (1973), 440—445.
28. STEIN, E. M. and WEISS, G. *Introduction to Fourier analysis on Euclidean spaces,* Princeton University Press, Princeton, (1971).
29. TRÈVES, F. *Topological vector spaces, distributions, and kernels,* Academic Press, New York, (1967).
30. YOSIDA, K. *Functional analysis,* 3rd ed.. Springer-Verlag, New York, (1971).
31. ZYGMUND, A. *Trigonometric series,* vol. II, Cambridge University Press, Cambridge, (1959).

Received September 26, 1974

G. B. Folland
Department of Mathematics
University of Washington
Seattle, Washington 98195
U.S.A.