HYPOELLIPTIC SECOND ORDER DIFFERENTIAL EQUATIONS

BY

LARS HÖRMANDER

The Institute for Advanced Study, Princeton, N.J., U.S.A.

1. Introduction

A linear differential operator P with C^∞ coefficients in an open set $\Omega \subset \mathbf{R}^n$ (or a manifold) is called hypoelliptic if for every distribution u in Ω we have

$$\operatorname{sing\,supp} u = \operatorname{sing\,supp} Pu,$$

that is, if u must be a C^∞ function in every open set where Pu is a C^∞ function. Necessary and sufficient conditions for P to be hypoelliptic have been known for quite some time when the coefficients are constant (see [3, Chap. IV]). It has also been shown that such equations remain hypoelliptic after a perturbation by a "weaker" operator with variable coefficients (see [3, Chap. VII]). Using pseudo-differential operators one can extend the class of admissible perturbations further; in particular one can obtain in that way many classes of hypoelliptic (differential) equations which are invariant under a change of variables (see [2]). Roughly speaking the sufficient condition for hypoellipticity given in [2] means that the differential equations with constant coefficients obtained by "freezing" the arguments in the coefficients at a point x shall be hypoelliptic and not vary too rapidly with x.

However, the sufficient conditions for hypoellipticity given in [2] are far from being necessary. For example, they are not satisfied by the equation

$$\frac{\partial^2 u}{\partial x^2} + x\frac{\partial u}{\partial y} - \frac{\partial u}{\partial t} = f, \tag{1.1}$$

for the operator obtained by freezing the coefficients at a point must operate along a two dimensional plane only so it cannot be hypoelliptic. But Kolmogorov [8] constructed already in 1934 an explicit fundamental solution of (1.1) which is a C^∞ function outside the diagonal, and this implies that (1.1) is hypoelliptic.

10 – 672909 *Acta mathematica* 119. Imprimé le 7 février 1968.

The arguments of Kolmogorov [8] can also be applied to the more general equation

$$-\frac{\partial u}{\partial x_0}+\sum_{j,k=1}^{n} a_{jk}\frac{\partial^2 u}{\partial x_j \partial x_k}+\sum_{j,k=1}^{n} b_{jk}x_j\frac{\partial u}{\partial x_k}+cu=f, \tag{1.2}$$

where a_{jk}, b_{jk} and c denote real constants, and the matrix $A=(a_{jk})$ is symmetric and positive semi-definite. If we take Fourier transforms in all variables except x_0 we are led to the equation

$$-\frac{\partial U}{\partial x_0}-A(\xi,\xi)\,U-\sum_{j,k=1}^{n} b_{jk}\frac{\partial(\xi_k U)}{\partial \xi_j}+cU=-\frac{\partial U}{\partial x_0}-(A(\xi,\xi)-c')\,U-\sum_{j,k=1}^{n} b_{jk}\xi_k\frac{\partial U}{\partial \xi_j}=F, \tag{1.3}$$

where $c'=c-\operatorname{Tr}B$. To obtain a fundamental solution of (1.2) with pole at (y_0, y) and vanishing when $x_0<y_0$ we wish to find a solution U of (1.3) when $x_0 \geqslant y_0$ such that $U=-e^{-i\langle y,\xi\rangle}$ when $x_0=y_0$. The characteristic equations for (1.3) are

$$dx_0=\frac{d\xi_j}{\sum b_{jk}\xi_k}=-\frac{dU}{U(A(\xi,\xi)-c')} \tag{1.4}$$

and have the solutions

$$x_0=y_0+t,\quad \xi(t)=(\exp Bt)\,\eta,\quad U=-\exp\left(-i\langle y,\eta\rangle-\int_0^t (A(\xi(s),\xi(s))-c')\,ds\right.$$

if we take the initial condition into account. Elimination of t and η gives

$$U(x_0,\xi)=-\exp\left(-i\langle y,(\exp B(y_0-x_0))\,\xi\rangle-\int_0^{x_0-y_0} A(\exp(-Bs)\,\xi,\exp(-Bs)\,\xi)\,ds\right.$$
$$\left.+(x_0-y_0)\,c'\right),\quad x_0>y_0;$$

and we set $U(x_0,\xi)=0$ when $x_0<y_0$. The quadratic form in the exponent is positive semi-definite, and it is positive definite unless for some $\xi \neq 0$ we have $A((\exp Bs)\,\xi, (\exp Bs)\,\xi)=0$ identically in s, that is, $A(B^k\xi, B^k\xi)=0$ for every k. This means that the null space of A contains a non-trivial invariant subspace for B. If this is not the case we obtain by inverting the Fourier transformation a two-sided fundamental solution which is a C^∞ function off the diagonal; for fixed x_0 and y_0 it is the exponential of a negative definite quadratic form in $(\exp {}^tBx_0)\,x-(\exp {}^tBy_0)\,y$ with eigenvalues $\to -\infty$ when $x_0-y_0 \to 0$. (The eigenvalues may have different orders of magnitude so the differentiability properties of the fundamental solution may be quite different in different directions, a typical feature of the subject of this paper.) Thus it follows that (1.2) *is hypoelliptic unless the null space of A contains a non-trivial subspace which is invariant for B.*

The results of Kolmogorov have been extended by Weber [11] and Il'in [4] to the equation

$$\sum_{i,j=1}^{n} a_{ij} \frac{\partial^2 u}{\partial y_i \partial y_j} + \sum_{1}^{n} a_i \frac{\partial u}{\partial y_i} + au + \sum_{1}^{m} b_i \frac{\partial u}{\partial x_i} - \frac{\partial u}{\partial t} = f, \tag{1.5}$$

where the coefficients are C^∞ functions of x, y, t, the matrix (a_{ij}) is positive definite and the matrix $(\partial b_i/\partial y_j)$ has rank m everywhere. The hypoellipticity follows from a construction of a fundamental solution by the E. E. Levi method starting from that given above after a change of variables to make $b_i = y_i$.

In this paper we shall give a nearly complete characterization of hypoelliptic second order differential operators P with real C^∞ coefficients. First it is easy to show, as we shall do in section 2, that the principal part must be semi-definite if P is hypoelliptic. In any open set where the rank is constant we can then write locally

$$P = \sum_{1}^{r} X_j^2 + X_0 + c, \tag{1.6}$$

where $X_0, \ldots, X_r$ denote first order homogeneous differential operators in an open set $\Omega \subset \mathbf{R}^n$ with C^∞ coefficients, and $c \in C^\infty(\Omega)$. We assume from now on that P has this form but do not necessarily require that the X_j are linearly independent at every point. There is of course a large freedom in the choice of the operators X_j. In particular, we may replace X_j by

$$X_j' = \sum_{1}^{r} c_{jk} X_k, \quad j = 1, \ldots, r,$$

where (c_{jk}) is an orthogonal matrix which is a C^∞ function of $x \in \Omega$; then X_0 is replaced by an operator X_0' such that $X_0' - X_0$ is a linear combination of $X_1, \ldots, X_r$ with C^∞ coefficients.

If the Lie algebra generated by $X_0, \ldots, X_r$ has constant rank $< n$ in a neighborhood of a point $x \in \Omega$, it follows from the Frobenius theorem that there exists a local change of variables near x so that P afterwards only acts in the variables $x_1, \ldots, x_{n-1}$. If the homogeneous equation $Pu = 0$ is satisfied by some non-trivial function it follows that P is not hypoelliptic, for a new solution is obtained by changing the definition of u to 0 on one side of a hyperplane $x_n = \text{constant}$. Thus the sufficient condition for hypoellipticity in the following theorem is essentially necessary also:

THEOREM 1.1. *Let P be written in the form* (1.6) *and assume that among the operators* $X_{j_1}, [X_{j_1}, X_{j_2}], [X_{j_1}, [X_{j_2}, X_{j_3}]], \ldots, [X_{j_1}, [X_{j_2}, [X_{j_3}, \ldots, X_{j_k}]]] \ldots$ *where* $j_i = 0, 1, \ldots, r$, *there exist n which are linearly independent at any given point in Ω. Then it follows that P is hypoelliptic.*

It is a simple exercise to verify that for the equation (1.2) the condition in Theorem 1.1 is the same that we needed to construct a smooth fundamental solution by the method of Kolmogorov [8].

As mentioned above we shall discuss necessary conditions for hypoellipticity in section 2. It is proved in section 3 that Theorem 1.1 is a consequence of certain *a priori* estimates and we make some preliminary steps toward proving these. However, they are proved completely only at the end of section 5 after a preliminary study of fractional differentiability of functions along a set of non-commuting vector fields has been made in section 4.

Finally we wish to mention that there is an extensive recent literature concerning global regularity of solutions of boundary problems for second order equations with semi-definite principal part. (See Kohn–Nirenberg [6, 7], Olejnik [10] and the references in these papers.) I am very much indebted to Professor Olejnik who first called my attention to the problem studied here and to Professor W. Feller who explained to me the probabilistic meaning of equations such as (1.1) and pointed out the existence of Kolmogorov's paper [8].

2. A necessary condition for hypoellipticity

A hypoelliptic differential equation with constant coefficients must have multiple characteristics if it is not elliptic (see [3, Theorem 4.17]). It is easy to extend this result to operators with variable coefficients, and this may have been done before. However, a proof will be given here since we do not know of any reference.

THEOREM 2.1. *Let $P(x, D)$ be a differential operator in $\Omega \subset \mathbf{R}^n$ with C^∞ coefficients, and let the principal symbol $p(x, \xi)$ be real. If for some $x \in \Omega$ one can find a real $\xi \neq 0$ such that*

$$p(x, \xi) = 0, \quad \textit{but} \quad \frac{\partial p(x, \xi)}{\partial \xi_j} \neq 0 \quad \textit{for some } j, \tag{2.1}$$

it follows that P is not hypoelliptic.

Proof. We may assume that the point x in (2.1) is 0. The classical integration theory for the characteristic equation (cf. [3, section 1.8] shows that there exists a real valued function φ in a neighborhood of 0 such that $\operatorname{grad} \varphi(0) = \xi$ and $p(x, \operatorname{grad} \varphi) = 0$. We can replace Ω by such a neighborhood of 0, and shall then determine a *formal* solution of the equation $P(x, D) u = 0$ by setting

$$u = \sum_0^\infty u_j t^{-j} e^{it\varphi},$$

where $u_j \in C^\infty(\Omega)$ and t is a parameter. Now we have

$$P(v\,e^{it\varphi}) = e^{it\varphi} \sum_0^m c_j t^j,$$

where $c_m = vp(x, \operatorname{grad} \varphi) = 0$ and $c_{m-1} = \sum_1^n A_j D_j v + Bv$ with $A_j = p^{(j)}(x, \operatorname{grad} \varphi)$. Hence we obtain

$$Pu = t^m \sum_0^\infty a_j t^{-j} e^{it\varphi},$$

where $$a_0 = 0, \quad a_1 = \sum_1^n A_j D_j u_0 + Bu_0, \quad a_k = \sum_1^n A_j D_j u_{k-1} + Bu_{k-1} + L_k,$$

where L_k is a linear combination of $u_0, \ldots, u_{k-2}$ and their derivatives. Since A_j are real and not all 0 we can if Ω is conveniently chosen successively find solutions $u_0, u_1, \ldots$ of these equations with $u_0(0) = 1$.

If the equation $Pu = 0$ were hypoelliptic we would have an *a priori* estimate

$$|\operatorname{grad} u(0)| \leqslant C\{\sup |u| + \sum_{|\alpha| \leqslant N} \sup |D^\alpha Pu|\}, \quad u \in C^\infty(\Omega), \tag{2.2}$$

where $\alpha = (\alpha_1, \ldots, \alpha_n)$ is a multiorder, $|\alpha| = \alpha_1 + \ldots + \alpha_n$, and $D^\alpha = (-i\partial/\partial x_1)^{\alpha_1} \ldots (-i\partial/dx_n)^{\alpha_n}$. Indeed, the set of all continuous and bounded functions in Ω with $D^\alpha Pu$ continuous and bounded in Ω for every α is then contained in C^∞ so (2.2) follows from the closed graph theorem. However, if we apply (2.2) to

$$\sum_0^{k-1} u_j t^{-j} e^{it\varphi},$$

where $N + m \leqslant k + 1$, it follows that the right-hand side of (2.2) is bounded when $t \to +\infty$ whereas the left-hand side is not. This proves the theorem.

COROLLARY 2.2. *For a second order hypoelliptic operator with real principal part, the principal part must be a semi-definite quadratic form.*

3. Preliminaries for the proof of Theorem 1.1

Let P be a differential operator of the form (1.6) where $X_j \in T(\Omega)$, the set of all homogeneous real first order differential operators in Ω with C^∞ coefficients, and $c \in C^\infty(\Omega)$. (We shall denote by $C^\infty(\Omega)$ the space of complex valued C^∞ functions in Ω and use the notation $C^\infty(\Omega, \mathbf{R})$ for the subset of real valued functions. Clearly $T(\Omega)$ is a $C^\infty(\Omega, R)$ module.) Alternatively we may of course regard $T(\Omega)$ as the space of C^∞ sections of the tangent bundle of Ω.

The starting point for the proof of Theorem 1.1 is an *a priori* estimate which is obtained by partial integration and also occurs frequently in the work of Kohn–Nirenberg [6, 7], Olejnik [10] and others. After noting that the adjoint of X_j is $-X_j+a_j$, where $a_j\in C^\infty(\Omega,\mathbf{R})$, the inequality is obtained by taking $v\in C_0^\infty(\Omega)$ and integrating by parts as follows:

$$\begin{aligned}-\operatorname{Re}\int v\,\overline{Pv}\,dx &= -\operatorname{Re}\sum_1^r\int v\,\overline{X_j^2v}\,dx-\operatorname{Re}\int v\,\overline{X_0v}\,dx-\int\operatorname{Re}c\,|v|^2\,dx\\ &=\operatorname{Re}\sum\int(X_jv-a_jv)\,\overline{X_jv}\,dx-\tfrac12\int X_0\,|v|^2\,dx-\int\operatorname{Re}c\,|v|^2\,dx\\ &=\sum\int|X_jv|^2\,dx+\int d\,|v|^2\,dx,\end{aligned}$$

where we have written

$$d=\tfrac12\sum(X_ja_j-a_j^2)\;-\tfrac12 a_0-\operatorname{Re}c.$$

Hence
$$\sum_1^r\|X_jv\|^2+\|v\|^2\leqslant C\|v\|^2-\operatorname{Re}\int v\,\overline{Pv}\,dx,\quad v\in C_0^\infty(K),\tag{3.1}$$

if K is a compact subset of Ω and $C_0^\infty(K)$ denotes the set of elements in $C^\infty(\Omega)$ with support in K. Here we have used the notation $\|\ \|$ for the L^2 norm.

For the left-hand side of (3.1) we introduce the notation

$$|||v|||^2=\sum_1^r\|X_jv\|^2+\|v\|^2.$$

To have a precise estimate for the right-hand side of (3.1) we also need the dual norm

$$|||f|||'=\sup_v\left|\int fv\,dx\right|/|||v|||,\quad v\in C_0^\infty(\Omega).$$

Then we have

$$-\operatorname{Re}\int v\,\overline{Pv}\,dx\leqslant|||v|||\;|||Pv|||'\leqslant\tfrac12(|||v|||^2+|||Pv|||'^2),$$

so with a new constant C we obtain from (3.1)

$$|||v|||^2\leqslant C\|v\|^2+|||Pv|||'^2,\quad v\in C_0^\infty(K).\tag{3.2}$$

Noting that
$$|||X_jf|||'\leqslant C\|f\|,\quad f\in C_0^\infty(K),\quad j=1,\dots,r,$$

we conclude that $|||X_j^2v|||'\leqslant C\|X_jv\|\leqslant C|||v|||$, $j=1,\dots,r$. Thus it follows from (3.2) that

$$|||v|||^2+|||X_0v|||'^2\leqslant C(\|v\|^2+|||Pv|||'^2),\quad v\in C_0^\infty(K).\tag{3.3}$$

Let $\|v\|_{(s)}$ denote the L^2 norm of the derivatives of v of order s (cf. [3, section 2.6]), defined by

$$\|v\|_{(s)}^2 = (2\pi)^{-n} \int |\hat{v}(\xi)|^2 (1+|\xi|^2)^s \, d\xi, \quad v \in C_0^\infty.$$

The main point of this paper is the proof given in sections 4 and 5 that for some $\varepsilon > 0$

$$\|v\|_{(\varepsilon)} \leqslant C(|||v||| + |||X_0 v|||'), \quad v \in C_0^\infty(K), \tag{3.4}$$

when the hypotheses of Theorem 1.1 are fulfilled. Combining this with (3.3) we obtain

$$\|v\|_{(\varepsilon)} \leqslant C(\|v\| + |||Pv|||'), \quad v \in C_0^\infty(K). \tag{3.5}$$

We shall now prove that it follows from (3.5) that P is hypoelliptic. The main step is the proof of the following proposition.

PROPOSITION 3.1. *Assume that* (3.5) *is valid for compact subsets* K *of* Ω. *Every* $v \in L^2(\Omega) \cap \mathcal{E}'(\Omega)$ *such that* $|||Pv|||' < \infty$ *is then in* $H_{(\varepsilon)}$.

We recall that $H_{(\varepsilon)}$ is the completion of C_0^∞ in the norm $\|\ \|_{(\varepsilon)}$ but refer to [3, section 2.6] for further discussion of this space.

Many statements of the same kind as Proposition 3.1 have been proved by Kohn and Nirenberg [6] but it seems that none of them contains Proposition 3.1 explicitly so we supply a proof here along lines similar to [3, Chap. VIII].

First note that (3.5) is valid for all $v \in H_{(2)}$ with compact support in Ω. Indeed, we can find a sequence $v_j \in C_0^\infty(K)$, where K is a compact neighborhood of $\operatorname{supp} v$, such that $D^\alpha v_j - D^\alpha v \to 0$, $j \to \infty$, when $j \leqslant 2$. Hence $\|Pv_j - Pv\| \to 0$, which implies that $|||Pv_j - Pv|||' \to 0$. In particular,

$$\overline{\lim}\, |||Pv_j|||' \leqslant |||Pv|||',$$

so it follows from (3.5) applied to v_j that

$$\overline{\lim}\, \|v_j\|_{(\varepsilon)} \leqslant C(\|v\| + |||Pv|||').$$

Hence (3.5) remains valid when $v \in H_{(2)}$ and $\operatorname{supp} v$ is in the interior of K.

If v satisfies the hypotheses of Proposition 3.1 we choose $\chi \in C_0^\infty(\Omega)$ so that $0 \leqslant \chi \leqslant 1$ and $\chi = 1$ in a neighborhood ω of $\operatorname{supp} v$, and we set

$$v_\delta = \chi(1 - \delta^2 \Delta)^{-1} v.$$

Here $(1-\delta^2\Delta)^{-1}v$ is defined as the inverse Fourier transform of $(1+\delta^2|\xi|^2)^{-1}\hat{v}(\xi)$. It is clear that v_δ is then in $H_{(2)}$, that $\operatorname{supp} v_\delta \subset \operatorname{supp} \chi \subset\subset \Omega$ and that $v_\delta \to v$ in L^2 norm when $\delta \to 0$. Hence we may apply (3.5) to v_δ and conclude that $\|v\|_{(\varepsilon)} < \infty$ provided that we can show that $|||Pv_\delta|||'$ remains bounded when $\delta \to 0$. This we shall do after a few simple remarks:

1°. The inverse Fourier transform K of $(1+|\xi|^2)^{-1}$ and all derivatives of K decrease exponentially at infinity. Since

$$(1-\delta^2\Delta)^{-1} v(x) = \delta^{-n} \int K((x-y)/\delta)\, v(y)\, dy$$

it follows that any derivative of $(1-\delta^2\Delta)^{-1} v(x)$ decreases faster than any power of δ when $\delta \to 0$ if $x \notin \omega$.

2°. If Q is a differential operator of order $j \leqslant 2$ with coefficients in C_0^∞, it follows that

$$\|(1-\delta^2\Delta)^{-1}\delta^j Q u\| \leqslant C\|u\|, \quad u \in L^2. \tag{3.6}$$

Indeed, the estimate $\|Q^*\delta^j(1-\delta^2\Delta)^{-1}u\| \leqslant C\|u\|$, $u \in C_0^\infty$, for the adjoint operator is trivial since $\delta^j(1+|\xi|)^j(1+\delta^2|\xi|^2)^{-1}$ is bounded.

3°. If Q is a differential operator of order $\leqslant 1$ with coefficients in C_0^∞, then

$$\|Q(1-\delta^2\Delta)^{-1}u - (1-\delta^2\Delta)^{-1}Qu\| \leqslant C\|u\|, \quad u \in C_0^\infty. \tag{3.7}$$

For writing $w = (1-\delta^2\Delta)^{-1}u$, we have $u = (1-\delta^2\Delta)w$ and

$$Qu = (1-\delta^2\Delta)\,Qw + \delta^2 Rw,$$

where $R = [\Delta, Q]$ is of second order. Multiplication by $(1-\delta^2\Delta)^{-1}$ gives

$$(1-\delta^2\Delta)^{-1}Qu - Q(1-\delta^2\Delta)^{-1}u = (1-\delta^2\Delta)^{-1}\delta^2 Rw,$$

and in view of 2° the L^2 norm of the right-hand side can be estimated by $\|w\| \leqslant \|u\|$.

4°. When $\chi_j \in C_0^\infty(\Omega)$ we have

$$|||\chi_1(1-\delta^2\Delta)^{-1}\chi_2 u||| \leqslant C|||u|||, \quad u \in C_0^\infty(\Omega), \qquad |||\chi_2(1-\delta^2\Delta)^{-1}\chi_1 f|||' \leqslant C|||f|||', \quad f \in \mathcal{D}'(\Omega).$$

The second inequality follows from the first which in turn is obvious since

$$\|X_j\chi_1(1-\delta^2\Delta)^{-1}\chi_2 u - (1-\delta^2\Delta)^{-1}X_j\chi_1\chi_2 u\| \leqslant C\|u\|$$

in view of 3°.

Proof of Proposition 3.1. We recall that with the notations introduced above we have to show that $|||Pv_\delta|||'$ is bounded when $\delta \to 0$. In the neighborhood ω of $\operatorname{supp} v$

we have $(1-\delta^2\Delta)\, v_\delta = v$. If we apply the operator P noting that $[X_j^2, \Delta] = X_j[X_j, \Delta] + [X_j, \Delta]\, X_j = 2X_j[X_j, \Delta] + [[X_j, \Delta], X_j]$, it follows that in ω

$$(1-\delta^2\Delta)\, Pv_\delta = Pv + \sum_1^r X_j \delta^2 B_j v_\delta + \delta^2 B_0 v_\delta,$$

where $B_0, \ldots, B_r$ are second order operators with compact support. In view of 1° it follows that we have everywhere

$$(1-\delta^2\Delta)\, Pv_\delta = Pv + \sum_1^r X_j \delta^2 B_j v_\delta + \delta^2 B_0 v_\delta + h_\delta,$$

where h_δ vanishes in ω, $\operatorname{supp} h_\delta \subset \operatorname{supp} \chi$ and $\|h_\delta\| \to 0$ when $\delta \to 0$. Hence

$$Pv_\delta = \chi_1 \{(1-\delta^2\Delta)^{-1} Pv + \sum_1^r (1-\delta^2\Delta)^{-1} X_j \delta^2 B_j v_\delta + (1-\delta^2\Delta)^{-1} \delta^2 B_0 v_\delta + (1-\delta^2\Delta)^{-1} h_\delta\},$$

where χ_1 is a function in $C_0^\infty(\Omega)$ which is equal to 1 in $\operatorname{supp}\chi$. Since $v = \chi v$, it follows from 4° that

$$|||\chi_1 (1-\delta^2\Delta)^{-1} Pv|||' \leq C\, |||Pv|||'.$$

The last two terms are bounded in L^2 norm in view of 2°, and since

$$(1-\delta^2\Delta)^{-1} X_j \delta^2 B_j v_\delta = X_j (1-\delta^2\Delta)^{-1} \delta^2 B_j v_\delta + [(1-\delta^2\Delta)^{-1}, X_j]\, \delta^2 B_j v_\delta,$$

we obtain using 3°

$$|||\chi_1 (1-\delta^2\Delta)^{-1} X_j \delta^2 B_j v_\delta|||' \leq C(\|(1-\delta^2\Delta)^{-1} \delta^2 B_j v_\delta\| + \|\delta^2 B_j v_\delta\|) \leq C'\, \|\delta^2 B_j v_\delta\| \leq C''\, \|v\|.$$

This completes the proof of Proposition 3.1.

PROPOSITION 3.2. *Assume that* (3.5) *is valid for compact subsets K of Ω. If $u \in \mathcal{D}'(\Omega)$ and $Pu = f \in H_{(s)}^{\mathrm{loc}}(\Omega)$, it follows that $u \in H_{(s+\varepsilon)}^{\mathrm{loc}}(\Omega)$. The same is true for open subsets of Ω, so in particular P is hypoelliptic.*

Proof. Since the statement is local we may assume that $u \in H_{(t)}^{\mathrm{loc}}(\Omega)$ for some t. It suffices to show that t can be replaced by $t+\varepsilon$ if $t \leq s$. Let E be a compactly supported pseudo-differential operator in Ω with symbol $e(\xi) = (1+|\xi|^2)^{t/2}$ (cf. [2]), and set $v = \chi E u$ where $\chi \in C_0^\infty(\Omega)$. If we can show that $v \in H_{(\varepsilon)}$ for every χ we will have $Eu \in H_{(\varepsilon)}^{\mathrm{loc}}$, hence $u \in H_{(t+\varepsilon)}^{\mathrm{loc}}$ since E is elliptic. It is clear that $v \in L^2(\Omega) \cap \mathcal{E}'(\Omega)$, so in view of Proposition 3.1 it only remains to show that $|||Pv|||' < \infty$. To do so we note that $\chi E f \in L^2$ since $t \leq s$, and form the difference

$$Pv - \chi E f = (P\chi E - \chi E P)\, u.$$

As in the proof of Proposition 3.1 we have $[X_j^2, \chi E] = 2X_j[X_j, \chi E] + [[X_j, \chi E], X_j]$ so it follows that

$$P\chi E - \chi E P = \sum_1^r X_j Q_j + Q_0,$$

where $Q_j = 2[X_j, \chi E]$ for $j = 1, \ldots, r$ and all Q_j are compactly supported pseudo-differential operators of order $\leqslant t$. Since $Q_j u \in L^2$ and has compact support, it follows that $|||P\chi Eu - \chi EPu|||' < \infty$, that is, $|||Pv|||' < \infty$. This completes the proof.

4. Differentiability along noncommuting vector fields

Let Ω be an open set in $\mathbf{R}^n$, K a compact subset of Ω, and let $X \in T(\Omega)$. We shall consider the one parameter (local) group of transformations in Ω defined by X. Thus let f be the solution of the initial value problem

$$\frac{df(x,t)}{dt} = X(f(x,t)), \quad f(x,0) = x. \tag{4.1}$$

It is clear that f is a C^∞ function from $K \times (-t_0, t_0)$ to Ω if t_0 is a small positive number depending on K and on X, and we have the group property

$$f(f(x,t), s) = f(x, t+s),$$

when $x \in K$ and $|t| + |s| < t_0$.

If u is a function in Ω we set

$$(e^{tX}u)(x) = u(f(x,t)).$$

When $|t| < t_0$ this defines a mapping from $C_0^\infty(K)$ to $C_0^\infty(\Omega)$ and one from $C^\infty(\Omega)$ to $C^\infty(K)$, and we have $e^{tX}e^{sX}u = e^{(t+s)X}u$ for small t and s. The differential equation for f gives

$$\frac{d(e^{tX}u)}{dt} = e^{tX}Xu.$$

The left-hand side is the limit of $(e^{(t+h)X}u - e^{tX}u)/h$ when $h \to 0$, hence also equal to $Xe^{tX}u$ by the same formula with u replaced by $e^{tX}u$ and t replaced by 0. Summing up, e^{tX} is a local one parameter group of transformations of functions in Ω, and

$$\frac{d(e^{tX}u)}{dt} = e^{tX}Xu = Xe^{tX}u.$$

When $u \in C^\infty$ we obtain the Taylor expansion at $t = 0$

$$e^{tX}u \sim \sum_0^\infty \frac{t^k X^k u}{k!}. \tag{4.2}$$

We shall be interested in Hölder continuity of functions along vector fields in the sense of L^2 norms. Thus we shall for $0<s\leqslant 1$ and $0<\varepsilon<t_0$ consider the norms

$$|u|^{\varepsilon}_{X,s}=\sup_{0<|t|<\varepsilon}\|e^{tX}u-u\|\,|t|^{-s},\quad u\in C_0^\infty(K), \tag{4.3}$$

where $\|\ \|$ denotes the L^2 norm. The norm (4.3) increases with ε, but since the difference between its values for two different choices of ε can be bounded by a constant times $\|u\|$, we shall usually omit ε from the notations below. An equivalent norm is of course obtained if we take $0<t<\varepsilon$. (Since our aim is to prove the *a priori* estimate (3.4) for $v\in C_0^\infty(K)$ we have chosen not to introduce the complete spaces corresponding to these norms and leave for the reader to state the implications for these spaces of the estimates proved below.)

LEMMA 4.1. *If $\varphi\in C^\infty(\Omega,\mathbf{R})$, it follows that*

$$|u|_{\varphi X,s}\leqslant C|u|_{X,s},\quad u\in C_0^\infty(K). \tag{4.4}$$

Proof. We keep the notation $f(x,t)$ used above so that $(e^{tX}u)(x)=u(f(x,t))$. Let $\tau(x,t)$ be the solution of the initial value problem

$$\frac{d\tau}{dt}=\varphi(f(x,\tau)),\quad \tau=0\quad\text{when}\quad t=0.$$

From the differential equation (4.1) we then obtain

$$\frac{df(x,\tau)}{dt}=(\varphi X)(f(x,\tau)),$$

Hence $e^{t\varphi X}u(x)=u(f(x,\tau))$, so that

$$\|e^{t\varphi X}u-u\|^2=\int|u(f(x,\tau(x,t)))-u(x)|^2\,dx.$$

Since τ depends on x we cannot compare this directly with (4.3), so we first note that for any σ

$$|u(f(x,\tau))-u(x)|^2\leqslant 2|u(f(x,\tau))-u(f(x,\sigma))|^2+2|u(f(x,\sigma))-u(x)|^2.$$

Integrating with respect to x and averaging over σ for $|\sigma|<|t|$, we obtain

$$\|e^{t\varphi X}u-u\|^2\leqslant|t|^{-1}\iint_{|\sigma|<|t|}|u(f(x,\tau))-u(f(x,\sigma))|^2\,dx\,d\sigma+2|u|^2_{X,s}|t|^{2s}.$$

In the integral we introduce new variables by setting

$$y = f(x, \sigma), \quad f(y, w) = f(x, \tau), \quad \text{that is,} \quad w + \sigma = \tau.$$

For fixed t and for $\sigma = 0$ we obtain

$$dy = dx + X\, d\sigma, \quad dw = d\tau - d\sigma.$$

When $t = 0$ we have $\tau = 0$, hence $d\tau = 0$, so for $t = \sigma = 0$ the Jacobian $D(y, w)/D(x, \sigma)$ is equal to -1. Hence it is arbitrarily close to -1 for sufficiently small σ and t. Since $\tau = O(t)$ we have $|w| \leqslant A|t|$ for some constant A when $|\sigma| \leqslant |t|$. Thus we conclude that for sufficiently small t

$$|t|^{-1} \iint_{|\sigma|<|t|} |u(f(x, \tau)) - u(f(x, \sigma))|^2\, dx\, d\sigma$$

$$\leqslant 2|t|^{-1} \iint_{|w|<A|t|} |u(f(y, w)) - u(y)|^2\, dy\, dw \leqslant 4A(A|t|)^{2s} |u|_{X,s}^2.$$

This completes the proof of (4.4).

We shall also use a universal s-norm defined by

$$|u|_s^\varepsilon = \sup_{|h|<\varepsilon} \|\tau_h u - u\|\, |h|^{-s},$$

where $(\tau_h u)(x) = u(x + h)$. If e_j is the field of unit vectors along the jth coordinate axis, we find immediately by using the triangle inequality that

$$|u|_s^\varepsilon \leqslant \sum_1^n |u|_{e_j, s}^\varepsilon. \tag{4.5}$$

On the other hand, we can estimate $|u|_{X,s}^\varepsilon$ by a constant times $|u|_s^\varepsilon$ for an arbitrary X. This is a special case (for $N = 1$) of the following

LEMMA 4.2. *Let $g(x, t)$ be a map from a neighborhood of $K \times 0$ in $\Omega \times \mathbf{R}$ to Ω such that $g(x, t) - x = O(t^N)$, $t \to 0$, where $N > 0$, and g is a C^∞ function of x which is continuous in t as well as its derivatives. Then we have for small $|t|$*

$$\int |u(g(x, t)) - u(x)|^2\, dx \leqslant C|t|^{2Ns} |u|_s^2, \quad u \in C_0^\infty(K). \tag{4.6}$$

Proof. The proof is parallel to that of Lemma 4.1. Thus we first compare with the translations τ_h and obtain

$$\int |u(g(x,t))-u(x)|^2\,dx \leqslant C|t|^{-Nn}\iint_{|h|<|t|^N} |u(g(x,t))-u(x+h)|^2\,dx\,dh + 2|u|_s^2|t|^{2Ns}.$$

We introduce new coordinates in the integral by setting $y=x+h$ and $y+w=g(x,t)$. For $t=0$ this is the linear transformation $y=x+h$, $w=-h$, which has determinant ± 1. For small t the Jacobian of the change of variables is therefore close to 1 in absolute value, and since $|w|<A|t|^N$ for some constant A in the domain of the new integral, the proof is concluded as that of Lemma 4.1.

Note that the norm $|\ |_s$ together with the L^2 norm is weaker than the usual s-norm $\|\ \|_{(s)}$ (cf. [3, section 2.6]) used in paragraph 3, but is stronger than the norm $\|\ \|_{(t)}$ when $t<s$.

If $X\in T(\Omega)$ we shall use the standard notation $\operatorname{ad} X$ for the differential operator from $T(\Omega)$ to $T(\Omega)$ defined by

$$(\operatorname{ad} X)\,Y=[X,Y],\quad Y\in T(\Omega).$$

Given elements $X_i\in T(\Omega)$, $i=0,\dots,r$, as in Theorem 1.1, and a multi-index I, that is, a sequence $(i_1,\dots,i_k)$ with $0\leqslant i_j\leqslant r$, we shall write

$$X_I=\operatorname{ad} X_{i_1}\dots\operatorname{ad} X_{i_{k-1}}X_{i_k}.$$

(Note the distinction between a multi-index and a multi-order as used in section 2.) We set $k=|I|$ and always assume that $|I|\neq 0$. The same notations will be used for other Lie algebras than $T(\Omega)$.

We can now state the main result to be proved in this section.

THEOREM 4.3. *Given $X_j\in T(\Omega)$ and $s_j\in(0,1]$, $j=0,\dots,r$, we denote by $T^s(\Omega)$ the $C^\infty(\Omega,\mathbf{R})$ submodule of $T(\Omega)$ generated by all X_I with $s(I)\geqslant s$, where*

$$\frac{1}{s(I)}=\sum_1^k\frac{1}{s_{i_j}}.$$

Assume that $T^s(\Omega)=T(\Omega)$ for some $s>0$. Then we have for every compact set $K\subset\Omega$ with C depending on X and K but not on u

$$|u|_{X,s}\leqslant C\left(\sum_0^r|u|_{X_j,s}+\|u\|\right),\quad u\in C_0^\infty(K),\quad X\in T^s(\Omega). \tag{4.7}$$

In particular,

$$|u|_s\leqslant C\left(\sum_0^r|u|_{X_j,s}+\|u\|\right),\quad u\in C_0^\infty(K),\quad \textit{if}\quad T^s(\Omega)=T(\Omega). \tag{4.8}$$

We remark that a slightly more precise version of a special case of (4.7) has been proved by Kohn [5] but his method does not seem applicable in the general case. The proof of (4.7) will be made by induction for increasing s starting from a point where we expect (4.8) to be valid. We begin with a simple lemma justifying that we have not considered more complicated commutators in Theorems 1.1 and 4.3.

LEMMA 4.4. *If* $1/t_1+1/t_2\leqslant 1/t_3$, *we have*

$$[T^{t_1}, T^{t_2}]\subset T^{t_3}.$$

Proof. Let I_1 and I_2 be multi-indices with $s(I_j)\geqslant t_j$ and let $\varphi_j\in C^\infty(\Omega)$. Then we have

$$[\varphi_1 X_{I_1}, \varphi_2 X_{I_2}]=\varphi_1(X_{I_1}\varphi_2)\,X_{I_2}-\varphi_2(X_{I_2}\varphi_1)\,X_{I_1}+\varphi_1\varphi_2[X_{I_1}, X_{I_2}].$$

Since $t_3\leqslant t_j$ for $j=1,2$, the first two terms on the right-hand side are in T^{t_3}, and so is the third since the Jacobi identity $\operatorname{ad}[X,Y]=[\operatorname{ad}X, \operatorname{ad}Y]$ gives that $[X_{I_1}, X_{I_2}]=(\operatorname{ad}X)_{I_1}X_{I_2}$ which written explicitly is a linear combination of elements X_J where $1/s(J)=1/s(I_1)+1/s(I_2)$.

We shall have to make repeated use of the Campbell–Hausdorff formula which can be stated as follows (cf. Hochschild [1], Chap. X): If x and y are two non-commuting indeterminates, we have in the sense of formal power series in x and y that $e^x e^y=e^z$ where

$$z=\sum_1^\infty(-1)^{n+1}n^{-1}\sum_{\alpha_i+\beta_i\neq 0}(\operatorname{ad}x)^{\alpha_1}(\operatorname{ad}y)^{\beta_1}\dots(\operatorname{ad}x)^{\alpha_n}(\operatorname{ad}y)^{\beta_n-1}y/c_{\alpha,\beta},$$

where $c_{\alpha,\beta}=\alpha!\,\beta!\,|\alpha+\beta|$. (When $\beta_n=0$ the term should be modified so that the last factor is $(\operatorname{ad}x)^{\alpha_n-1}x$.) The important facts for us are that the terms of order 1 are $x+y$, that those of order two are $\frac{1}{2}[x,y]$, and that all terms of higher order are (repeated) commutators of x and y.

We shall use the Campbell–Hausdorff formula to derive a product decomposition of e^{x+y}. With the notations used above we have

$$e^{-y}e^{-x}e^{x+y}=e^{-z}e^{x+y}=e^{r_2},$$

where $r_2=-z+x+y+\frac{1}{2}[-z,x+y]+\dots=-\frac{1}{2}[x,y]+\dots$, the dots indicating terms of order at least three which are linear combinations of commutators. Writing $z_2=-\frac{1}{2}[x,y]$, we form $e^{-z_2}e^{r_2}=e^{r_3}$. The Campbell–Hausdorff formula gives $r_3=z_3+\dots$ where z_3 is a linear combination of commutators of x and of y of degree three, and the dots indicate a formal series whose terms are commutators of degree at least four. Proceeding in this way, we choose for every integer $k \geq 2$ a linear combination z_k of commutators of x and y of order k such that

$$e^{-z_k} e^{r_k} = e^{r_{k+1}},$$

where r_k is a formal series whose terms are commutators of x and y of order at least k. Thus we have

$$e^{-z_k} e^{-z_{k-1}} \dots e^{-z_2} e^{-y} e^{-x} e^{x+y} = e^{r_{k+1}}, \tag{4.9}$$

so that e^{x+y} is to a high degree of accuracy approximated by the product

$$e^x e^y e^{z_2} \dots e^{z_k}.$$

LEMMA 4.5. *Let $X, Y \in T(\Omega)$ and denote by Z_j the linear combination of commutators of j factors X and Y obtained by replacing x and y by X and Y above. Let $0 < \sigma \leq 1$, and let N be an integer ≥ 2. Then we have for small t and $u \in C_0^\infty(K)$*

$$\| e^{t(X+Y)} u - u \| \leq C(\| e^{tX} u - u \| + \| e^{tY} u - u \| + \sum_2^{N-1} \| e^{t^j Z_j} u - u \| + t^{\sigma N} |u|_\sigma). \tag{4.10}$$

Proof. The operator

$$H_N^t = \exp(-t^{N-1} Z_{N-1}) \dots \exp(-t^2 Z_2) \exp(-tY) \exp(-tX) \exp t(X+Y)$$

is induced by a mapping in Ω since every factor is. Hence there exists a C^∞ function $h_N(x,t)$ from a neighborhood of $K \times 0$ to Ω such that $H_N^t v(x) = v(h_N(x,t))$. From (4.9) and (4.2) it follows that $H_N^t v - v = O(t^N)$ if $v \in C^\infty$. Taking for v a coordinate function we conclude that

$$h_N(x,t) - x = O(t^N),$$

so Lemma 4.2 gives $$\| H_N^t v - v \| \leq C |t|^{\sigma N} |v|_\sigma.$$

Now we have for any bounded operators $S_1, \dots, S_k$ in L^2

$$\| S_1 \dots S_k u - u \| = \| \sum_{j=1}^k S_1 \dots S_{j-1}(S_j u - u) \| \leq \sum_{j=1}^k \| S_1 \| \dots \| S_{j-1} \| \, \| S_j u - u \|. \tag{4.11}$$

Since $$\exp t(X+Y) = \exp(tX) \exp(tY) \exp(t^2 Z_2) \dots \exp(t^{N-1} Z_{N-1}) H_N^t$$

and the norm of each factor is bounded uniformly in t the inequality (4.10) follows.

Lemmas 4.1 and 4.5 together will allow us to prove (4.7) for arbitrary $X \in T^s(\Omega)$ after the estimate has been established for a set of generators. We shall therefore study next some identities which give control of the commutators of the given operators X_j.

Let $x_1, \ldots, x_k$ be k non-commuting indeterminates. By the Campbell–Hausdorff formula we have

$$e^{x_{k-1}} e^{x_k} e^{-x_{k-1}} e^{-x_k} = e^{z_{k-1}},$$

where $z_{k-1} = [x_{k-1}, x_k] + \ldots$, the dots indicating a formal series all terms of which are commutators of at least three factors equal to x_k or x_{k-1}; obviously both x_k and x_{k-1} must occur at least once in every one of them. We now form successive formal power series $z_{k-1}, z_{k-2}, \ldots, z_1$ by setting

$$e^{x_j} e^{z_{j+1}} e^{-x_j} e^{-z_{j+1}} = e^{z_j}, \quad j = 1, \ldots, k-2.$$

Then e^{z_1} is a product of n_k factors $e^{\pm x_j}$, where $n_2 = 4$ and $n_{k+1} = 2 + 2n_k$, that is, $n_k = 3 \cdot 2^{k-1} - 2$, and $z_1 = c + \ldots$ where $c = \operatorname{ad} x_1 \operatorname{ad} x_2 \ldots \operatorname{ad} x_{k-1} x_k$, the dots denoting a series with terms of higher order, each of which is a commutator containing each x_j at least once. As in the discussion preceding Lemma 4.5 we can use the Campbell–Hausdorff formula again to show that for any N we can write

$$e^c = e^{z_1} e^{c_1} e^{c_2} \ldots e^{c_\nu} e^r,$$

where each c_j is a commutator formed from $x_1, \ldots, x_k$ which contains each x_j at least once and some x_j twice, and r is a formal sum of commutators of at least N factors x_j. If we recall the definition of e^{z_1} we have thus found an approximate representation of e^c by products of $e^{\pm x_j}$ and e^{c_j} where c_j are commutators of higher order than c. This allows us to prove the final lemma needed for the proof of Theorem 4.3.

LEMMA 4.6. *Given X_j and s_j, $j = 0, \ldots, r$, as in Theorem 4.3, we set $m_j = 1/s_j$ and $m(I) = 1/s(I)$ when I is a multi-index. Let $\sigma > 0$. Then we have for small $t > 0$ and an arbitrary multi-index I*

$$\|\exp(t^{m(I)} X_I) u - u\| \leqslant C_1 t \sum_0^r |u|_{X_j, s_j} + C_2 t |u|_\sigma, \quad u \in C_0^\infty(K), \tag{4.12}$$

where C_1 and C_2 are constants and C_1 only depends on r and σ, not on $X_0, \ldots, X_r, s_0, \ldots, s_r$.

Proof. Since $m_j \geqslant 1$, we have $m(I) \geqslant |I|$, so (4.12) follows from Lemma 4.2 with $C_1 = 0$ if $\sigma |I| \geqslant 1$. If N is an integer with $N\sigma \geqslant 1$, we may thus prove (4.12) by induction for decreasing $|I|$, starting when $|I| = N$.

Replacing the indeterminates x_j in the discussion preceding Lemma 4.6 by $t^{m_{i_j}} X_{i_j}$, we obtain as in the proof of Lemma 4.5 an identity

$$\exp(t^{m(I)}X_I)=\prod \exp(\pm t^{m_j}X_j)\exp(t^{m(I_1)}X_{I_1})\dots\exp(t^{m(I_\nu)}X_{I_\nu})H_N^t, \tag{4.13}$$

where the product contains $3\cdot 2^{|I|-1}-2$ factors as described above, the multi-indices $I_1,\dots,I_\nu$ have greater length than I and

$$H_N^t v(x)=v(h_N(x,t))$$

with a C^∞ function $h_N(x,t)$ of x, depending continuously on t, such that $h_N(x,t)-x=O(t^N)$, $t\to 0$. From Lemma 4.2 we obtain

$$\|H_N^t u-u\|\leqslant C\,t^{N\sigma}|u|_\sigma,\quad u\in C_0^\infty(K).$$

In view of (4.11) if follows that for small t

$$\|\exp(t^{m(I)}X_I)u-u\|\leqslant 2^{|I|+1}\sum_0^r\|\exp(t^{m_j}X_j)u-u\|$$
$$+2\sum_1^\nu\|\exp(t^{m(I_j)}X_{I_j})u-u\|+C\,t^{\sigma N}|u|_\sigma,$$

for the norm of each factor in (4.13) is close to 1 for small t. We can apply the inductive hypothesis to the terms in the second sum, and since $\sigma N\geqslant 1$, the estimate (4.12) follows.

Remark. Since C_1 does not depend on the choice of $X_0,\dots,X_r$, it follows that for multi-indices I containing some index $\geqslant 1$, we have for every $\varepsilon>0$

$$\|\exp(t^{m(I)}X_I)u-u\|\leqslant\varepsilon t|u|_{X_0,s_0}+C_\varepsilon t\left(\sum_1^r|u|_{X_j,s_j}+|u|_\sigma\right),\quad u\in C_0^\infty(K). \tag{4.12'}$$

In fact, X_I does not change if we replace X_0 by εX_0 provided that at the same time we replace X_j by $\varepsilon^{-\gamma}X_j$ for a suitable $\gamma>0$ when $j\geqslant 1$.

Proof of Theorem 4.3. Choose $\sigma>0$ so that $T^\tau(\Omega)=T(\Omega)$ for some $\tau>\sigma$. We wish to prove that

$$|u|_{X,s}\leqslant C_X\left(\sum_0^r|u|_{X_j,s_j}+|u|_\sigma\right),\ u\in C_0^\infty(K),\ X\in T^s(\Omega). \tag{4.14}$$

This estimate is trivial if $s\leqslant\sigma$ and it follows from Lemma 4.6 if X is any one of the commutators X_I which generate $T^s(\Omega)$. In view of Lemma 4.1 the estimate (4.14) remains valid for $X=\varphi X_I$ if $\varphi\in C^\infty(\Omega)$. By definition every $X\in T^s$ is therefore a finite sum of vector fields for which (4.14) is valid. If X, $Y\in T^s(\Omega)$ are vector fields for which (4.14) is valid, we apply Lemma 4.5 with $\sigma N\geqslant s$ noting that if follows from Lemma 4.4 that $Z_j \in T^{s/j}$. Assuming as we may that (4.14) has already been proved when s is replaced by a number $\leqslant s/2$, we conclude that (4.14) is valid when X is replaced by $X+Y$. Hence (4.14) follows.

Now recall that $T^{\tau}(\Omega) = T(\Omega)$ for some $\tau > \sigma$. Thus we obtain from (4.14) the estimate

$$|u|_\tau \leqslant C\left(\sum_0^r |u|_{X_j, s_j} + |u|_\sigma\right), \quad u \in C_0^\infty(K), \tag{4.15}$$

if we take (4.5) into account. Since $\tau > \sigma$ we have for any $\delta > 0$

$$|u|_\sigma \leqslant \delta |u|_\tau + C_\delta \|u\|. \tag{4.16}$$

If we combine (4.15) and (4.16) taking $\delta C < \frac{1}{2}$, we conclude that

$$|u|_\sigma \leqslant |u|_\tau \leqslant C'\left(\sum_0^r |u|_{X_j, s_j} + \|u\|\right), \quad u \in C_0^\infty(K).$$

Using this estimate in the right-hand side of (4.14) we have proved (4.7).

5. Smoothing and estimates

In section 3 we have proved that Theorem 1.1 is a consequence of the *a priori* estimate (3.4). We recall that

$$|||v|||^2 = \sum_1^r \|X_j v\|^2 + \|v\|^2,$$

so the right-hand side of (3.4) gives us control of $|v|_{X_j, 1}$ when $j = 1, \ldots, r$. However, the information given about $X_0 v$ is in a weaker norm which prevents us from applying Theorem 4.3. To study the differentiability of v in the direction X_0 we consider $f(t) = \|e^{tX_0} v - v\|$. Differentiation with respect to t gives, if v is real as we may well assume

$$df(t)^2/dt = 2(e^{tX_0} X_0 v, e^{tX_0} v - v).$$

Let us assume for a moment that e^{tX_0} preserves the norms $|||\ |||$ and $|||\ |||'$, although we shall see below that this is far from true. Then we would obtain

$$df(t)^2/dt \leqslant 4 |||X_0 v|||' \ |||v|||,$$

hence

$$f(t) \leqslant t^{\frac{1}{2}} (|||v||| + |||X_0 v|||').$$

Thus we would have control of $|v|_{X_0, \frac{1}{2}}$ and could apply Theorem 4.3 with $s_0 = \frac{1}{2}$, $s_1 = \ldots = s_r = 1$, and this would give (3.4).

To examine the validity of the preceding arguments we must consider $|||e^{tX_0}v|||$, thus the L^2 norm of $X_j e^{tX_0} v$ for $j=1, \ldots, r$. This is essentially the same as the L^2 norm of $e^{-tX_0} X_j e^{tX_0} v$. Now

$$e^{-tX_0} X_j e^{tX_0} = e^{-t \operatorname{ad} X_0} X_j = \sum_0^\infty (-t)^k (\operatorname{ad} X_0)^k X_j / k!,$$

where the first equality is a definition motivated by the second one which means that the two sides have the same Taylor expansion in t. This follows immediately from the fact that left and right multiplication by X_0 commute and that by definition their difference is ad X_0. Since we have no information about the differentiability of v in the direction $(\operatorname{ad} X_0)^k X_j$ when $k \neq 0$, the argument as given above breaks down. However, we note that the derivative in this direction occurs with a factor t^k, which indicates that we can impose sufficient smoothness on v by a regularization which does not change v too much for small t. This we shall do in the following discussion which aims at proving that $f(t^2)/t$ can be bounded by the right-hand side of (3.4). It is in fact permissible to allow in the right-hand side of the estimates any quantities which by the results of section 4 can be estimated by a small constant times $|v|_{X_0, \frac{1}{2}}$ and a large constant times $|||v|||$.

The first step is to study regularization along a vector field $X \in T(\Omega)$. Let $u \in C_0^\infty(K)$, $K \subset\subset \Omega$, and assume that $e^{\tau X}$ maps $C_0^\infty(K)$ into $C_0^\infty(\Omega)$ for $|\tau| \leqslant 1$. With $\varphi \in C_0^\infty(-1, 1)$ we set

$$\varphi_X u = \int e^{\tau X} u \varphi(\tau)\, d\tau.$$

This operator is smoothing in the direction X, for

$$X \varphi_X u = \int \frac{d}{d\tau} e^{\tau X} u \varphi(\tau)\, d\tau = -\varphi'_X u = \int (u - e^{\tau X} u)\, \varphi'(\tau)\, d\tau.$$

It follows that
$$\| X \varphi_X u \| \leqslant \int |d\varphi| \sup_{|\tau|<1} \| e^{\tau X} u - u \|, \tag{5.1}$$

$$\| \varphi_X u - u \| \leqslant \sup_{|\tau|<1} \| e^{\tau X} u - u \| \quad \text{if} \quad \varphi \geqslant 0 \quad \text{and} \quad \int \varphi\, dx = 1. \tag{5.2}$$

We shall later have to consider the commutator of φ_X with operators $Y \in T(\Omega)$, so we note the formulas

$$Y \varphi_X u = \int e^{\tau X} (e^{-\tau \operatorname{ad} X} Y)\, u \varphi(\tau)\, d\tau, \tag{5.3}$$

$$\varphi_X Y u = \int (e^{\tau \operatorname{ad} X} Y)\, e^{\tau X} u \varphi(\tau)\, d\tau. \tag{5.4}$$

Each term in the Taylor expansion of $e^{\tau \operatorname{ad} X}$ will thus give an analogous expression with a smoothing operator defined by some other function and acting on the other side of another differential operator.

Besides the quite specific smoothing along certain vector fields using the operators φ_X we shall also employ the usual smoothing in all variables. Thus let $\Phi \in C_0^\infty(B)$, where B is the unit ball in $\mathbf{R}^n$, and set

$$\Phi_\varepsilon u(x) = \int u(x - \varepsilon h)\, \Phi(h)\, dh.$$

Then we obtain

$$\varepsilon \| D_j \Phi_\varepsilon u \| \leq \int |D_j \Phi|\, dx \sup_{|h|<1} \| u(x - \varepsilon h) - u(x) \|, \tag{5.1$'$}$$

$$\| \Phi_\varepsilon u - u \| \leq \sup_{|h|<1} \| u(x - \varepsilon h) - u(x) \| \quad \text{if} \quad \Phi \geq 0 \quad \text{and} \quad \int \Phi\, dx = 1. \tag{5.2$'$}$$

Instead of (5.3) and (5.4) we shall use Friedrichs's lemma ([3, Theorem 2.4.3]) which gives for every $Y \in T(\Omega)$

$$\| (Y \Phi_\varepsilon - \Phi_\varepsilon Y) u \| \leq C \| u \|, \quad u \in C_0^\infty(K), \tag{5.5}$$

where C is uniformly bounded for small ε if Y lies in a bounded set in $T(\Omega)$.

As in section 4 the notation I will stand for a multi-index and X_I for the corresponding commutator. We set $s_0 = \frac{1}{2}$, $s_1 = \ldots = s_r = 1$ and define $s(I)$ and $m(I) = 1/s(I)$ as in Theorem 4.3 and Lemma 4.6. Thus $m(I)$ is the sum of the length $|I|$ of I and the number of indices in I which are equal to 0.

Let σ be a positive number chosen so small that with the notations of Theorem 4.3 we have $T^s(\Omega) = T(\Omega)$ for some $s > \sigma$. As in section 4 we shall allow $|u|_\sigma$ to occur in the right-hand side of our estimates and use it to take care of various remainder terms in Taylor expansions. Let $\mathfrak{J}$ denote the set of all I with $\sigma m(I) \leq 1$ and $|I| < m(I) < 2|I|$; the latter condition means that I shall contain indices equal to 0 as well as indices $\neq 0$. Set

$$M(u) = |||u||| + |||X_0 u|||' + \sum_{I \in \mathfrak{J}} |u|_{X_I, s(I)} + |u|_\sigma.$$

Our aim is to show that

$$|u|_{X_0, \frac{1}{2}} \leq C M(u), \quad u \in C_0^\infty(K). \tag{5.6}$$

By (4.12)′ we can estimate $|u|_{X_I, s(I)}$ by a small constant times $|u|_{X_0, \frac{1}{2}}$ and a large constant times $|||u||| + |u|_\sigma$ when $I \in \mathfrak{J}$. Hence (5.6) implies

$$\sum_0^r |u|_{X_j, s_j} + \|u\| \leqslant C' \left(|||u||| + |||X_0 u|||' + |u|_\sigma \right), \quad u \in C_0^\infty(K). \tag{5.7}$$

Let $s > \sigma$ but $T^s(\Omega) = T(\Omega)$. Then it follows from (5.7) and Theorem 4.3 that

$$|u|_s \leqslant C'' \left(|||u||| + |||X_0 u|||' + |u|_\sigma \right), \quad u \in C_0^\infty(K),$$

and since $|u|_\sigma \leqslant \delta |u|_s + C_\delta \|u\|$ for any $\delta > 0$, we obtain with another constant C

$$|u|_s \leqslant C \left(|||u||| + |||X_0 u|||' \right), \quad u \in C_0^\infty(K),$$

hence

$$\sum_0^r |u|_{X_j, s_j} \leqslant C \left(|||u||| + |||X_0 u|||' \right), \quad u \in C_0^\infty(K). \tag{5.8}$$

In view of Theorem 4.3 we conclude

THEOREM 5.1. *Let $X_0, \dots, X_r$ satisfy the hypotheses of Theorem 4.3 and set $s_0 = \frac{1}{2}$, $s_1 = \dots = s_r = 1$. Then we have*

$$|u|_{X, s} \leqslant C(X) \left(|||u||| + |||X_0 u|||' \right), \quad u \in C_0^\infty(K), \quad X \in T^s(\Omega), \tag{5.9}$$

where $|||\ |||$ and $|||\ |||'$ are defined in section 3.

Clearly Theorem 5.1 completes the proof of Theorem 1.1, so all that remains now is to prove (5.6).

We give $\mathfrak{J}$ a total ordering so that $m(I)$ is an increasing function of $I \in \mathfrak{J}$ and set

$$S_t u = \prod_{I \in \mathfrak{J}} \varphi_{t^{m(I)} X_I} \Phi_{t^{1/\sigma}} u,$$

where φ and Φ are functions satisfying the hypotheses of (5.2) and (5.2)′. The factors in the product are taken from left to right in increasing order of I. If $J \in \mathfrak{J}$ we shall also write $S_t^J u$ for the similar expression with the product restricted to all I with $I \geqslant J$. We also set $\mathfrak{J}' = \mathfrak{J} \cup \infty$ and $S_t^\infty = \Phi_{t^{1/\sigma}}$, and define $\infty > I$ for every $I \in \mathfrak{J}$.

If follows immediately from (5.2), (5.2)′, the definition of the norms and (4.11) that

$$\|S_t u - u\| \leqslant C\, t\, M(u). \tag{5.10}$$

We want to estimate $\|e^{t^2 X_0} u - u\|$. Noting that

$$e^{t^2 X_0} u - u = e^{t^2 X_0}(u - S_t u) + e^{t^2 X_0} S_t u - S_t u + S_t u - u$$

and that the norm of $e^{t^2 X_0}$ as an operator in L^2 is uniformly bounded, we conclude that

$$\| e^{t^2 X_0} u - u \| \leqslant C t M(u) + \| e^{t^2 X_0} S_t u - S_t u \|. \tag{5.11}$$

The advantage of this is that the regularity built into $S_t u$ will make it possible to estimate the last term by applying the argument outlined at the beginning of the section but which was then merely heuristic. We need the following lemma which shows how differentiability is successively introduced by the regularizations in S_t. (Note that $S_t = S_t^J$ when J is the smallest element in $\mathfrak{J}$.)

LEMMA 5.2. *For every $J \in \mathfrak{J}'$ we have for small $t > 0$*

$$\sum_1^n \| t^{1/\sigma} D_j S_t^J u \| \leqslant C t M(u), \tag{5.12}$$

$$\sum_{J \leqslant I \in \mathfrak{J}} \| t^{m(I)} X_I S_t^J u \| \leqslant C t M(u), \tag{5.13}$$

$$\sum_0^r \| [t^{m_j} X_j, S_t^J] u \| \leqslant C t M(u). \tag{5.14}$$

Proof. For $J = \infty$ the estimate (5.12) follows from (5.1)$'$. As a superposition of compositions with C^∞ maps, each factor in the operators S_t^J is uniformly bounded in the $H_{(s)}$ norm for every s, so (5.12) is valid for all $J \in \mathfrak{J}'$. For $J = \infty$ the statement (5.13) is void and (5.14) is very much weaker than (5.5). When proving the lemma we may therefore assume that it has already been proved for larger J and arbitrary $\varphi \in C_0^\infty(-1, 1)$. In the proof of (5.13) we must then distinguish between two different cases:

1°. $I > J$. Let J' be the smallest element in $\mathfrak{J}'$ larger than J; then $I \geqslant J'$. We shall use (5.3) with Y replaced by $t^{m(I)} X_I$ and X replaced $t^{m(J)} X_J$. This allows us to let X_I pass through the first regularizer and we obtain

$$t^{m(I)} X_I S_t^J u = \int e^{\tau t^{m(J)} X_J} \left\{ \sum_{\nu < N} (\mathrm{ad} - \tau t^{m(J)} X_J)^\nu t^{m(I)} X_I S_t^{J'} A u / \nu ! + t^{N m(J) + m(I)} Y_{t,\tau} S_t^{J'} u \right\} \varphi(\tau)\, d\tau,$$

where $Y_{t,\tau}$ belongs to a bounded set when $t \to 0$. If $N m(J) + m(I) \geqslant 1/\sigma$ if follows from (5.12) that we have the desired estimate for the term involving the remainder term $Y_{t,\tau}$. Since $(\mathrm{ad}\ X_J)^\nu X_I = X_{I'}$ for some I' with $I' \geqslant I > J$ we have $I' \geqslant J'$, so the inductive hypothesis concerning (5.13) allows us to estimate the terms in the sum to the extent that they are not taken care of by (5.12).

2°. $I=J$. With J' defined as above we have in view of (5.1)

$$\|t^{m(J)}X_J S_t^J u\| \leqslant C \sup_{|\tau|\leqslant 1} \| e^{\tau t^{m(J)}X_J} S_t^{J'} u - S_t^{J'} u\|.$$

The proof of (5.10) applies without change to prove that the right-hand side can be bounded by $CtM(u)$.

It remains to prove (5.14). We can write

$$[t^{m_j}X_j, S_t^J]\, u = \varphi_{t^{m(J)}X_J}[t^{m_j}X_j, S_t^{J'}]\, u + [t^{m_j}X_j, \varphi_{t^{m(J)}X_J}]\, S_t^{J'}\, u.$$

By the inductive hypothesis concerning (5.14) it suffices to consider the second term. Now (5.3) gives

$$[t^{m_j}X_j, \varphi_{t^{m(J)}X_J}]\, v = \int e^{\tau t^{m(J)}}X_J \left(\sum_{0<\nu<N} (\mathrm{ad} - \tau t^{m(J)}X_J)^\nu\, t^{m_j}X_j v/\nu! + t^{NM(J)+m_j}Y_{t,\tau}v\right) \varphi(\tau)\, d\tau$$

with $Y_{t,\tau}$ as above in 1°. When $\nu \neq 0$ we have $(\mathrm{ad}-X_J)^\nu X_j = X_{I'}$ for some $I' \in \mathfrak{J}'$ with $I' \geqslant J'$ or else $\sigma m(I) \geqslant 1$, so we obtain (5.14) from the inductive hypotheses concerning (5.12) and (5.13).

Proof of (5.6). Our aim is to estimate the right-hand side of (5.11) so we introduce for $0 \leqslant \tau \leqslant t^2$ the function

$$f(\tau) = \| e^{\tau X_0} S_t u - S_t u\|.$$

Differentiation of f^2 gives

$$\begin{aligned} f(\tau) f'(\tau) &= (e^{\tau X_0} X_0 S_t u, e^{\tau X_0} S_t u - S_t u) \\ &= (e^{\tau X_0}[X_0, S_t]\, u, e^{\tau X_0} S_t u - u) + (e^{\tau X_0} S_t X_0 u, e^{\tau X_0} S_t u - S_t u). \end{aligned}$$

Using the Cauchy–Schwarz inequality and (5.14), we obtain

$$f(\tau) f'(\tau) \leqslant Ct^{-1} M(u) f(\tau) + (X_0 u, (e^{\tau X_0} S_t)^* (e^{\tau X_0} S_t u - S_t u)). \tag{5.15}$$

We shall prove that the last expression can be estimated by $CM(u)^2$. Admitting this estimate for a moment, we obtain the integral inequality

$$ff' \leqslant C\left(\frac{f}{t} + M\right) M \quad \text{when} \quad 0 < \tau < t^2; \;\; f(0) = 0.$$

If $g = f/Mt$, this reduces to $gg' \leqslant C(g+1)\, t^{-2}$, so

$$\int_0^g g\,dg/(g+1) \leqslant C\tau t^{-2} \leqslant C \quad \text{when} \quad \tau \leqslant t^2.$$

This implies that $g \leqslant C'$ when $\tau \leqslant t^2$, hence $f(t^2) \leqslant C'M(u)t$, and (5.6) follows in view of (5.11).

What remains is therefore to show that

$$|||(e^{\tau X_0}S_t)^*(e^{\tau X_0}S_t u - S_t u)||| \leqslant CM(u). \tag{5.16}$$

Write $$N_t(v) = \sum_1^r \|X_j v\| + \sum_{I \in \mathfrak{J}} \|t^{m(I)-1} X_I v\| + t^{1/\sigma - 1}\sum_1^n \|D_j v\| + \|v\|.$$

Then we have $$N_t(S_t u) \leqslant CM(u). \tag{5.17}$$

This follows immediately from Lemma 5.2 if we note that $X_j S_t u = S_t X_j u + [X_j, S_t]u$ and recall again that S_t has uniformly bounded norm in L^2. The norm N_t is quite well behaved under translations; indeed, we shall prove that

$$N_t(e^{\tau X_J} v) \leqslant CN_t(v), \; 0 < \tau \leqslant t^{m(J)}, \tag{5.18}$$

provided that the multi-index J contains 0.

To prove (5.18) we let $Y = X_j$, $j = 1, \dots, r$, or $Y = X_I$ where $I \in \mathfrak{J}$, and note that

$$Y e^{\tau X_J} S_t u = e^{\tau X_J}(e^{-\operatorname{ad} \tau X_J} Y) S_t u.$$

Now $$\|(\operatorname{ad} \tau X_J)^k Y v\| \leqslant \|(\operatorname{ad} t^{m(J)} X_J)^k Y v\|$$

when $0 \leqslant \tau \leqslant t^{m(J)}$, so we obtain the desired bound for each term in the Taylor expansion of $e^{-\operatorname{ad}\tau X_J}$. The error term can be estimated by using the last sum in the definition of N_t, so (5.18) follows.

In particular, we obtain from (5.17) and (5.18)

$$N_t(e^{\tau X_0} S_t u - S_t u) \leqslant CM(u) \quad \text{when} \quad 0 \leqslant \tau \leqslant t^2. \tag{5.19}$$

Now the adjoint of a translation e^Y is equal to $J_Y e^{-Y}$ where for Y in a suitable neighborhood of 0 the Jacobian J_Y has a uniform bound together with as many derivatives as we wish. In view of (5.18) it follows that the adjoint of $e^{\tau X_0}$ for $0 \leqslant \tau \leqslant t^2$ and of $e^{\tau X_I}$ for $0 \leqslant \tau \leqslant t^{m(I)}, I \in \mathfrak{J}$, are uniformly bounded with respect to the norms N_t, as is the adjoint of the operator $\Phi_{t^{1/\sigma}}$. Since $(e^{\tau X_0}S_t)^*$ for $0 \leqslant \tau \leqslant t^2$ is a superposition with finite total mass of operators which have uniformly bounded norm with respect to N_t, it follows from (5.19) that

$$N_t((e^{\tau X_0}S_t)^*(e^{\tau X_0}S_t u - S_t u)) \leqslant CM(u),$$

and this implies (5.16). Thus we have completed the proof of (5.6) and so we have proved Theorems 5.1 and 1.1.

References

[1]. Hochschild, G., *The structure of Lie groups*. Holden–Day Inc., San Francisco, London, Amsterdam, 1965.

[2]. Hörmander, L., Pseudo-differential operators and hypoelliptic equations. To appear in *Amer. Math. Soc. Proc. Symp. Pure Math.*, 10 (1967).

[3]. ——, *Linear partial differential operators*. Springer-Verlag, Berlin–Göttingen–Heidelberg, 1963.

[4]. Il'in, A. M., On a class of ultraparabolic equations. (Russian.) *Doklady Akad. Nauk SSSR*, 159 (1964), 1214–1217. Also in *Soviet Math. Dokl.*, 5 (1964), 1673–1676.

[5]. Kohn, J. J., Boundaries of complex manifolds. *Proc. Conf. Complex Analysis (Minneapolis 1964)*, 81–94. Springer Verlag, Berlin, 1965.

[6]. Kohn, J. J. & Nirenberg, L., Non coercive boundary value problems. *Comm. Pure Appl. Math.*, 18 (1965), 443–492.

[7]. ——, Degenerate elliptic-parabolic equations of second order. To appear in *Comm. Pure Appl. Math.*

[8]. Kolmogorov, A. N., Zufällige Bewegungen. *Ann. of Math.* (2), 35 (1934), 116–117.

[9]. Nirenberg, L. & Trèves, F., Solvability of a first order linear partial differential equation. *Comm. Pure Appl. Math.*, 16 (1963), 331–351.

[10]. Olejnik, O. A., Linear second order equations with non-negative characteristic form. (Russian.) *Mat. Sb.*, 69 (1966), 111–140.

[11]. Weber, M., The fundamental solution of a degenerate partial differential equation of parabolic type. *Trans. Amer. Math. Soc.*, 71 (1951), 24–37.

Received June 12, 1967