SUBELLIPTIC EIGENVALUE PROBLEMS

C. Fefferman
Princeton University

D.H. Phong
Columbia University

This article studies subellipticity and asymptotic eigenvalue distribution of second-order partial differential operators with non-negative principal symbols. We obtain definitive results for self-adjoint operators with real coefficients. Deep theorems on subellipticity and related topics have been obtained by Hörmander [6], Kohn [7], and Rothschild and Stein [10] for sums of squares of vector fields; and by Olenik and Radkevitch [9] for general second-order equations; while subelliptic eigenvalue asymptotics were studied by Menikoff and Sjöstrand [8]. The interest of our work is that it gives sharp necessary and sufficient conditions for operators not assumed to be written as sums of squares. It would be interesting to understand also non-self-adjoint equations and equations with complex coefficients.

To fix the notation, let L be a second-order self-adjoint differential operator on a compact manifold M with smooth measure μ. Assume that in local coordinates

$$L = -\sum_{ij} a^{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_j b_j(x)\frac{\partial}{\partial x_j} + c(x),$$

NSF Grant No. MCS80-03072.

NSF Grant No. MCS78-27119.

where a^{ij}, b_j, c are real and $(a^{ij}(x))$ is positive semidefinite. We want to understand the following:

(A) When does L satisfy the subelliptic estimates

$$\operatorname{Re}\langle Lu, u\rangle + C\|u\|^2 \geq c\|u\|^2_{(\varepsilon)} \quad \text{for } u \in C^\infty(M); \tag{1}$$

$$\|Lu\| + C\|u\| \geq c\|u\|_{(2\varepsilon)} \quad \text{for } u \in C^\infty(M)? \tag{2}$$

(B) What is the asymptotic behavior as $\lambda \to \infty$ of $N(\lambda, L)$, the number of eigenvalues of L which are $< \lambda$?

These questions are answered in terms of a family of "non-Euclidean balls," which we now define.

A tangent vector $X = \sum_j \gamma_j \frac{\partial}{\partial x_j}$ at $x \in M$ is said to be *subunit* for L if $(\gamma_j \gamma_k) \leq (a^{jk}(x))$ as matrices. One easily checks that this notion is independent of the particular coordinate chart. For $x \in M$ and $\rho > 0$, the "ball" $B_L(x, \rho)$ consists of all the points $y \in M$ that can be joined to x by a Lipschitz path $\gamma : [0, \rho] \to M$ for which $\frac{d}{dt}\gamma(t)$ is a subunit vector for L at $\gamma(t)$ for almost every t.

By $B_E(x, \rho)$ we denote an ordinary Euclidean ball of radius ρ about x. Note that if $-L$ is the Laplacian for a metric ds^2 on M, then $B_L(x, \rho)$ agrees with the usual ball of radius ρ in ds^2. See also Rothschild and Stein [10], in which suitable non-Euclidean balls play an important role.

Theorem 1. For L as above, the estimates (1) *and* (2) *are equivalent, and they both hold if and only if*

$$B_E(x, \rho) \subseteq B_L(x, C\rho^\varepsilon) \text{ for } x \in M, \rho > 0.$$

Theorem 2. Assume L satisfies (1) *and* (2). *Then for large λ, $N(\lambda, L)$ is bounded above and below by constant multiples of*

$$\tilde{N}(\lambda, L) = \int_M \frac{d\mu(x)}{\mu(B_L(x, \lambda^{-1/2}))}.$$

The rest of this paper sketches the proofs of Theorems 1 and 2. We make essential use of results and techniques in [2] and [5], with which we assume the reader is familiar. The new tools needed in our proofs are the following geometric lemmas.

Lemma 1. $B_E(x, \lambda) \subseteq B_L(x, C\lambda^{\varepsilon})$ *if and only if*

$$B_E(x, \lambda) \subseteq B_{L-\lambda^{2N}\Delta}(x, C\lambda^{\varepsilon}),$$

provided $N \geq \varepsilon^{-50}$. *Moreover, these inclusions imply that*

$$\mu(B_L(x, \lambda)) \sim \mu(B_{L-\lambda^{2N}\Delta}(x, \lambda)).$$

Lemma 2. In R^{n+1}, *set*

$$L_\lambda = -\left(\frac{\partial}{\partial t}\right)^2 - \sum_{ij} a^{ij}(t, x)\frac{\partial^2}{\partial x_i \partial x_j} - \lambda^{2N}\Delta_x;$$

in R^n, *set*

$$\bar{L}_\lambda = -\sum_{ij} \overline{a^{ij}}(x)\frac{\partial^2}{\partial x_i \partial x_j} - \lambda^{2N}\Delta_x,$$

where

$$\overline{a^{ij}}(x) = \lambda^{-1}\int_{|t| \leq \lambda} a^{ij}(t, x)\,dt.$$

Then the non-Euclidean balls B_{L_λ}, $B_{\bar{L}_\lambda}$ *about the origin are related by* $B_{L_\lambda}(c\lambda) \subseteq \{|t| \leq \lambda\} \times B_{\bar{L}_\lambda}(\lambda) \subseteq B_{L_\lambda}(C\lambda)$.

Lemma 1 and the terms $\lambda^{2N}\Delta_x$ in Lemma 2 are technicalities; while Lemma 2 contains new geometric information, which is essential both here and in our paper [4].

To prove Lemma 2, we need the following result, which forms an L^∞-analogue of the spectral decomposition theorem in [5].

For

$$L = -\sum_{ij} a^{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \cdots$$

defined on the unit cube Q^0, and for $u \in L^\infty(Q^0)$, define

$$L_K(u) = \inf_{v \in C^1(Q^0)} \left\{ K\|u - v\|^2 + \left\| \sum_{ij} a^{ij}(x) \frac{\partial v}{\partial x_i} \frac{\partial v}{\partial x_j} \right\| + K^{-2N} \|\nabla u\|^2 \right\},$$

where the norms are taken in $L^\infty(Q^0)$.

Lemma 3. $c_N (L + L')_K(u) \leq L_K(u) + L'_K(u)$.

We first sketch how Lemmas 1 and 2 yield Theorems 1 and 2 by using [2 and 5]. Next we prove Lemma 1 and reduce Lemma 2 to Lemma 3. Finally, we sketch the proof of Lemma 3.

Proof of Theorem 1. It is easy to show that (1) and (2) are equivalent. In fact, (2) asserts that $(L + CI)^2 > \Lambda^{4\varepsilon}$ as self-adjoint operators on $L^2(M)$. One knows from operator theory that $A^2 \geq B^2$ implies $A \geq B$ for positive operators A, B, so $L + CI \geq \Lambda^{2\varepsilon}$, which is (1). On the other hand, assume (1), and note that $\mathrm{Re}\langle Lu, \Lambda^{2\varepsilon}u\rangle \leq c_1\|u\|^2_{(2\varepsilon)} + 10c_1^{-1}\|Lu\|^2$ for a small c_1 to be picked later. However, $\mathrm{Re}\langle Lu, \Lambda^{2\varepsilon}u\rangle = \mathrm{Re}\langle L\Lambda^{\varepsilon}u, \Lambda^{\varepsilon}u\rangle + \mathrm{Re}\langle \Lambda^{\varepsilon}[\Lambda^{\varepsilon}, L]u, u\rangle$. One computes that $\Lambda^{\varepsilon}[\Lambda^{\varepsilon}, L] = T_1 + T_2$ with $T_1 \in S^{2\varepsilon+1}$ skew-adjoint and $T_2 \in S^{2\varepsilon}$. Therefore,

$$|\mathrm{Re}\langle \Lambda^{\varepsilon}[\Lambda^{\varepsilon}, L]u, u\rangle| = |\mathrm{Re}\langle T_2 u, u\rangle| \leq c_1\|u\|^2_{(2\varepsilon)} + 10c_1^{-1}\|u\|^2.$$

Moreover, our assumption (1) implies

$$\mathrm{Re}\langle L\Lambda^{\varepsilon}u, \Lambda^{\varepsilon}u\rangle \geq c\|u\|^2_{(2\varepsilon)} - C\|u\|^2_{(\varepsilon)} \geq \frac{1}{2}c\|u\|^2_{(2\varepsilon)} - C'\|u\|^2.$$

Putting these estimates together, we obtain

$$\frac{c}{2}\|u\|^2_{(2\varepsilon)} - C'\|u\|^2 - c_1\|u\|^2_{(2\varepsilon)} - 10c_1^{-1}\|u\|^2 \leq c_1\|u\|^2_{(2\varepsilon)} + 10c_1^{-1}\|Lu\|^2,$$

which implies (2) if we pick $c_1 < \frac{c}{10}$. We are indebted to J. J. Kohn and E. M. Stein for stimulating conversations on estimate (2).

The hard part of Theorem 1 is to show that (1) is equivalent to $B_E(x, \lambda) \leq B_L(x, C\lambda^{\varepsilon})$. To prove this, we formulate a stronger result, which can be proved by induction on the dimension.

Given a second-order operator L defined on R^n, and large constants K, S, we say that $L \geq K$ *microlocally in* $|x| < 1$, $|\xi| \sim S$

if $\theta(x, D)^*(L - KI)\theta(x, D) \geq -CI$ for some symbol $\theta \in S^0$ satisfying $\theta(x, \xi) = 1$ for $|x| \leq 1$, $|\xi| \sim S$.

One checks easily that (1) amounts to saying that in local coordinates $L \geq (\text{const})K$ microlocally in $|x| \leq 1$, $|\xi| \sim S$, with $K = S^{2\varepsilon}$.

Therefore, Theorem 1 will follow at once from Lemma 1 and the

Main Estimate. Assume $K \geq S^{\delta}$ *and* $N \geq N_{\min}(\delta)$. *Then*

(i) $L \geq (\text{const})K$ *microlocally in* $|x| \leq 1$, $|\xi| \sim S$ *iff*

(ii) $B_E(x, (\text{const})S^{-1}) \subseteq B_{L-K^{-N}\Delta}(x, K^{-1/2})$.

So out task is to prove the main estimate. In one variable, the result is trivial; in n dimensions we proceed by induction on n as follows.

We first reduce matters to the case

$$L = -\left(\frac{\partial}{\partial t}\right)^2 + \tilde{L}\left(t, y, \frac{\partial}{\partial y}\right).$$

This is done by localizing. We make a Calderón-Zygmund decomposition of $\{|x| \leq 1\}$ into cubes $\{Q_\nu\}$ of diameters δ_ν, stopping at Q_ν when $\max_{ij} \max_{x \in Q_\nu^*} |a^{ij}(x)| \geq C\delta_\nu^2$. (Recall

$$L = -\sum_{ij} a^{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \cdots .)$$

We may assume all the $\delta_\nu \geq S^{-1}$, since otherwise (i) and (ii) are both false. In each Q_ν, make a change of coordinates from $x \in Q_\nu^*$ to $(t, y) \in$ unit cube, so that L goes over to

$$L_\nu = -\left(\frac{\partial}{\partial t}\right)^2 + \tilde{L}_\nu\left(t, y, \frac{\partial}{\partial y}\right).$$

Now the sumbolic calculus of pseudodifferential operators shows that $L \geq (\text{const})K$ microlocally for $|\xi| \sim S$ if and only if each $L_\nu \geq (\text{const})K$ microlocally for $|(\tau, \eta)| \sim S\delta_\nu$. Moreover, under the coordinate change $x \to (t, y)$, $B_E(x, S^{-1})$ goes over to $B_E((t, y), (S\delta_\nu)^{-1})$, while $B_{L + \text{junk}}(x, K^{-1/2})$ goes over to $B_{L + \text{junk}}((t, y), K^{-1/2})$. Thus, the main estimate holds for L if and only if it holds for L_ν. So it is enough to look at

$$L = -\left(\frac{\partial}{\partial t}\right)^2 + L(t, y, \partial y).$$

Also, we may assume $K \leq S^2$, since otherwise (i) and (ii) are again obviously false. Now, however, Lemma 2 shows that in place of $B_{L - K^{-N}\Delta}((\bar{t}, \bar{y}), K^{-1/2})$, we may take

$$\{|t - \bar{t}| \leq K^{-1/2}\} \times B_{\bar{L}_{\bar{t}} - K^{-N}\Delta}(\bar{y}, K^{-1/2})$$

in (ii). Here

$$\bar{L}_{\bar{t}}(y, \partial y) = K^{1/2}\int_{|t-\bar{t}| \leq K^{-1/2}} \tilde{L}(t, y, \partial_y)\,dt.$$

Also, the results in Fefferman and Phong [5] show that (i) is equivalent to saying that $\bar{L}_{\bar{t}} \geq$ (const)K microlocally in $|y| \leq 1$, $|\eta| \sim S$ for each $\bar{t}$. Therefore, our main estimate holds for L, provided it holds for $\bar{L}_{\bar{t}}$ for every $\bar{t}$. Since $\bar{L}_{\bar{t}}$ is an operator in $n - 1$ variables, the induction step is complete. So we know both the main estimate and Theorem 1.

Proof of Theorem 2. Using symbolic calculus and the minimax formula for eigenvalues, one can easily localize the problem by assuming L to be elliptic outside $\{|x| \leq 1\}$. Another application of symbolic calculus localizes the problem still further to the operators L_ν in the proof of Theorem 1. So we may assume

$$L = -\left(\frac{\partial}{\partial t}\right)^2 + \tilde{L}(t, y, \partial y)$$

with L elliptic outside $\{|t|, |y| \leq 1\}$. Since L is subelliptic, one checks that $N(\lambda, L)$ is comparable to $N(\lambda; L - \lambda^{-2N}\Delta)$ for large N depending on the ε in (1). So we may also assume that

$$\tilde{L}(t, y, \partial_y) = -\sum_{ij} a^{ij}(t, y)\frac{\partial^2}{\partial y_j\,\partial y_k} + \cdots,$$

with $(a^{ij}(t, y)) \geq \lambda^{-2N}(\delta_{ij})$. (This uses Lemma 1.) Now define

$$\bar{L}_{\bar{t}}(y, \partial_y) = \lambda^{+1/2}\int_{|t-\bar{t}|\leq\lambda^{+1/2}} \tilde{L}(t, y, \partial_y)dt.$$

Lemma 2 shows that $B_L((t, y), \lambda^{1/2})$ is comparable to $\{|t - \bar{t}| \leq \lambda^{-1/2}\} \times B_{\bar{L}_{\bar{t}}}(y, \lambda^{-1/2})$, while the results of Fefferman and Phong [2] show that $N(\lambda, L)$ is comparable to

$$\lambda^{+1/2}\int_{|\bar{t}|\leq 1} N(\lambda, \bar{L}_{\bar{t}})d\bar{t}$$

for large λ. Therefore, the formula

$$N(\lambda, L) \sim \tilde{N}(\lambda, L) = \int_M \frac{d\mu(x)}{\mu(B_L(x, \lambda^{-1/2}))}$$

is easily deduced from the corresponding formula for the $\bar{L}_{\bar{t}}$. Since $\bar{L}_{\bar{t}}$ is an operator in $n - 1$ variables, we can now proceed by induction on the dimension.

The above argument shows that N and $\tilde{N}$ are comparable in the sense that $c\tilde{N}(c\lambda, L) \leq N(\lambda, L) \leq C\tilde{N}(C\lambda, L)$. A final application of Lemma 2 gives $\tilde{N}(C\lambda, L) \leq C'\tilde{N}(\lambda, L)$ and $\tilde{N}(c\lambda, L) \geq c'N(\lambda, L)$, so that $c''\tilde{N}(\lambda, L) \leq N(\lambda, L) \leq C''\tilde{N}(\lambda, L)$ for large λ.

Proof of Lemma 1. Assuming $B_E(x, \lambda^{N/2}) \subseteq B_{L-\lambda^{2N}\Delta}(x, \lambda)$ for $x \in M$, $\lambda > 0$, we shall prove that

$$(+) \qquad B_{L-\lambda^{2N}\Delta}(x, \lambda) \subseteq B_L(x, C\lambda).$$

This easily implies Lemma 1. In fact, using $C\lambda^{\varepsilon}$ in place of λ and εN in place of N in (+), we obtain $B_{L-\lambda^{2N}\Delta}(x, C\lambda^{\varepsilon}) \subseteq B_L(x, C'\lambda^{\varepsilon})$, so that $B_E(\ldots) \subseteq B_{L-\lambda^{2N}\Delta}(\ldots)$ implies $B_E(\ldots) \subseteq B_L(\ldots)$. The converse implication is trivial, since $B_L(x, C\lambda^{\varepsilon}) \subseteq B_{L-\lambda^{2N}\Delta}(x, C\lambda^{\varepsilon})$. Inclusion (+) also yields the part of Lemma 1 on $\mu(B_L(x, \lambda))$, since $\mu(B_L(x, \lambda)) \leq \mu(B_{L-\lambda^{2N}\Delta}(x, \lambda)) \leq \mu(B_L(x, C\lambda))$. This implies $\mu(B_{L-c\lambda^{2N}\Delta}(x, c\lambda)) \leq \mu(B_L(x, \lambda)) \leq \mu(B_{L-\lambda^{2N}\Delta}(x, C\lambda))$.

An application of Lemma 2 shows that the terms on the extreme left and right are comparable, so that indeed (+) implies all of Lemma 1.

We proceed to prove (+).

Given a point $y \in B_{L-\lambda^{2N}\Delta}(x, \lambda)$, we may join x to y by a geodesic $\gamma : [0, \lambda] \to M$ in the metric

$$ds^2 = \sum_{ij} g_{ij}(x)\, dx_i\, dx_j,$$

where

$$(g_{ij})^{-1} = (a^{ij} + \lambda^{2N}\delta_{ij}) > 0,\ L = -\sum_{ij} a^{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \cdots .$$

Thus $\gamma(0) = x$, $\gamma(\lambda) = y$, and $\frac{d}{dt}\gamma(t)$ is a subunit vector for $L - \lambda^{2N}\Delta$.

We shall perturb $\gamma(t)$ to a broken path $\gamma_\#(t)$ that starts at x and ends very near y, so that $\frac{c\,d\gamma_\#(t)}{dt}$ is a subunit vector for L.

We construct $\gamma_\#$ using the following elementary remarks.

(a) PERTURBATION OF TANGENT VECTORS

Say that $X = \sum_j \gamma_j \frac{\partial}{\partial x_j}$ is a subunit vector for $L - \lambda^{2N}\Delta$ at x^0, and suppose $|x' - x^0| < c_1\lambda^N$. Then there is a tangent vector $Y = \sum_j \gamma_j' \frac{\partial}{\partial x_j}$ at x' satisfying

(i) $|\gamma_j - \gamma_j'| \leq C\lambda^N$,

(ii) cY is a subunit vector for L at x',

(iii) $\gamma_i' = \sum_j a^{ij}(x')\xi_j'$ with $|\xi'| \leq C\lambda^{-N}$.

For our hypothesis on X is that

$$\left(\sum_j \gamma_1 \eta_1\right)^2 \leq \sum_{ij}(a^{ij}(x^0) + \lambda^{2N}\delta_{ij})\eta_i\eta_j$$

for $\eta \in R^n$. Applying Lemma 4.1 in [5] to the function

$$f(t) = |\eta|^{-2}\sum_{ij}(a^{ij}(x' + t(x^0 - x')) + \lambda^{2N}\delta_{ij})\eta_i\eta_j \geq 0$$

yields

$$\sum_{ij} (a^{ij}(x^0) + \lambda^{2N}\delta_{ij})\eta_i\eta_j \leq C\sum_{ij} (a^{ij}(x') + \lambda^{2N}\delta_{ij})\eta_i\eta_j + C|x' - x^0|^2|\eta|^2,$$

so that for $|x^0 - x'| \leq c_1\lambda^N$, we have

$$c\left(\sum_j \gamma_j\eta_j\right)^2 \leq \sum_{ij} (a^{ij}(x') + \lambda^{2N}\delta_{ij})\eta_i\eta_j.$$

Thus $c\tilde{X}$ is a subunit vector for $L - \lambda^{2N}\Delta$ at x', where $\tilde{X} = \sum_j \gamma_j \frac{\partial}{\partial x_j}$ at x'. Next we rotate the coordinate axes so that $a^{ij}(x')$ is diagonalized: $(a_{ij}(x')) = (\lambda_i\delta_{ij})$, $\lambda_i \geq 0$. Note that conclusions (i), (ii), and (iii) are unaffected by the rotation. Now we know that $c\tilde{X}$ is a subunit for $L - \lambda^{2N}\Delta$, that is,

$$\sum_i \frac{\gamma_i^2}{\lambda_i + \lambda^{2N}} \leq C.$$

Define $Y = \sum_i \gamma_i' \frac{\partial}{\partial x_i}$ at x', where $\gamma_i' = \gamma_i$ if $\lambda_i \geq \lambda^{2N}$, $\gamma_i' = 0$ if $\lambda_i < \lambda^{2N}$. Setting $\xi_i' = \gamma^i/\lambda_i$ for $\lambda_i \geq \lambda^{2N}$, $\xi_i' = 0$ for $\lambda_i < \lambda^{2N}$, we see easily that (i), (ii), and (iii) hold.

(b) ESTIMATES FOR SECOND DERIVATIVES OF HAMILTONIAN PATHS

Say $H = \frac{1}{2}\sum_{ij} a^{ij}(x)\xi_i\xi_j$ and initially $|\xi| \leq C\lambda^{-N}$. Hamiltonian's equations

$$\dot{x}^i = \sum_j a^{ij}(x)\xi_j, \qquad \dot{\xi}_1 = -\sum_{ij} \frac{\partial a^{ij}(x)}{\partial x_1}\xi_i\xi_j,$$

$$\ddot{x}^i = \sum_j a^{ij}(x)\dot{\xi}_j + \sum_{j\ell} \frac{\partial a^{ij}(x)}{\partial x_\ell}\dot{x}_\ell\xi_j$$

imply $|\dot{x}| \leq C|\xi|$, $|\dot{\xi}| \leq C|\xi|^2$, $|\ddot{x}| \leq C|\xi|^2$. So if we flow for time $\Delta t \leq c_1\lambda^N$, then $|\dot{\xi}| \leq C'\lambda^{-N}$, $|\ddot{x}| \leq C''\lambda^{-2N}$ along the path.

Now we can construct the broken path $\gamma_\#(t)$ mentioned above. Suppose we have constructed $\gamma_\#(t)$ for $0 \leq t \leq \tau_K$, and that $\gamma_\#(0) = \gamma(0)$, $|\gamma(t) - \gamma_\#(t)| \leq C_+\lambda^N\tau_K$ for $0 \leq t \leq \tau_K$. (The large constant C_+ will be picked later. Note that the assertion is vacuous for $\tau_0 = 0$.) Apply (a) above with $x^0 = \gamma(\tau_K)$, $x' = \gamma_\#(\tau_K)$, and

$X = \frac{d}{dt}\gamma(t)\Big|_{\tau_\kappa}$. Since λ is small, we have

$$|x^0 - x'| \leq C_+\lambda^N\tau_\kappa \leq C_+\lambda^{N+1},$$

so the hypotheses of (a) are satisfied. With Y, ξ' as in (a), we now define $\gamma_\#(t)$ for $\tau_\kappa \leq t \leq \tau_{\kappa+1}$ as the projection onto the x-coordinate of the Hamiltonian curve for

$$H = \frac{1}{2}\sum_{ij} a^{ij}(x)\xi_i\xi_j$$

starting at (x', ξ') for $t = \tau_\kappa$. Since $\frac{d}{dt}\gamma_\#\Big|_{\tau_\kappa} = Y$, and since H is conserved along the path, it follows from (a)(ii) above that $c\frac{d}{dt}\gamma_\#$ is a subunit vector for L throughout $\tau_\kappa \leq t \leq \tau_{\kappa+1}$. The estimates of (b) show that $\left|\frac{d^2}{dt^2}\gamma_\#\right| \leq C\lambda^{-2N}$ for $\tau_\kappa \leq t \leq \tau_{\kappa+1}$, provided $\tau_{\kappa+1} - \tau_\kappa < c_1\lambda^N$. Similarly, $\left|\frac{d^2}{dt^2}\gamma\right| \leq C\lambda^{-2N}$, while $\left|\frac{d\gamma}{dt}(\tau_\kappa) - \frac{d\gamma_\#}{dt}(\tau_\kappa)\right| = |\overline{X} - \overline{Y}| \leq C\lambda^N$ by (a)(i). So if $\tau_{\kappa+1} - \tau_\kappa = \lambda^{3N}$, then we have for $\tau_\kappa \leq t \leq \tau_{\kappa+1}$ that

$$\left|\frac{d\gamma}{dt} - \frac{d\gamma_\#}{dt}\right| \leq C\lambda^{+N} + C\lambda^{-2N}|t - \tau_\kappa| \leq C'\lambda^N.$$

Our assumption $|\gamma(\tau_\kappa) - \gamma_\#(\tau_\kappa)| \leq C_+\lambda^N\tau_\kappa$ now implies for $\tau_\kappa \leq t \leq \tau_{\kappa+1}$ that

$$|\gamma(t) - \gamma_\#(t)| \leq C_+\lambda^N\tau_\kappa + C'\lambda^N(t - \tau_\kappa) \leq C_+\lambda^N\tau_{\kappa+1},$$

provided we pick $C_+ \geq C'$. Our construction of $\gamma_\#(t)$ is complete, with $\tau_\kappa = \kappa\cdot\lambda^{3N}$.

The proof of (+) is now easy. Given $y \in B_{L-\lambda^{2N}\Delta}(x, \lambda)$ and γ as above, we constructed a broken curve $\gamma_\# : [0, \lambda] \to M$ so that $c\frac{d}{dt}\gamma_\#$ is subunit for L, while $\gamma_\#(0) = x$, $|\gamma_\#(\lambda) - y| \leq C_+\lambda^{N+1}$. Thus $y_\# = \gamma_\#(\lambda) \in B_L(x, C\lambda)$, while $y \in B_E(y_\#, \lambda^N) \subseteq B_{L-(\lambda^2)^{2N}\Delta}(y^\#, \lambda^2)$. Repeating the process yields a sequence of paths $\gamma_\#^\mu : [t_\mu, t_{\mu+1}] \to M$ with tangent vectors subunit for L, so that $\gamma_\#^0$ joins $x = \gamma_\#^0(t_0 = 0)$ to $y_\#^1 = \gamma_\#(t_1)$, $\gamma_\#^\mu$ joins $y_\#^\mu$ to $y_\#^{\mu+1}$, $y_\#^\mu \to y$, and $\sum_\mu t_\mu \leq C'\lambda$.

Combining the $\gamma_{\#}^{\mu}$ into a single Lipschitz path $\gamma_{\#\#}$, we see that $y \in B_L(x, C'\lambda)$ as required.

Proof of Lemma 2. Since $(a^{ij}(t, x)) \leq C(\overline{a^{ij}}(x) + \lambda^{2N}\delta_{ij})$ for $|t| \leq \lambda$, we see at once that $B_{L_\lambda}(c\lambda) \subseteq B_{L_\lambda^*}(\lambda) \subseteq \{|t| \leq \lambda\} \times B_{\overline{L}_\lambda}(\lambda)$, where $L_\lambda^* = (\partial/\partial t)^2 + \overline{L}_\lambda(y, \lambda y)$. So our problem is to show that $\{|t| \leq \lambda\} \times B_{\overline{L}_\lambda}(\lambda) \subseteq B_{L_\lambda}(C\lambda)$. To see this, let $u(t, y)$ be the distance from (t, y) to the origin induced by L_λ. Note that $u_\lambda \in \mathrm{Lip}(1)$, since $(a^{ij} + \lambda^{2N}\delta_{ij}) > 0$. Since $\partial/\partial t$ is a subunit vector for L_λ, we have $t^{-2}|u_\lambda(t, x) - u_\lambda(0, x)|^2 \leq 1$, while also

$$\sum_{ij}(a^{ij}(t, y) + \lambda^{2N}\delta_{ij})\frac{\partial u_\lambda(t, y)}{\partial y_i}\frac{\partial u_\lambda(t, y)}{\partial y_j} \leq 1,$$

where u_λ is differentiable. By mollifying $u_\lambda(t, \cdot)$ slightly, we obtain $v = v_{\lambda, t} \in C^\infty$, so that

$$t^{-2}\|u_\lambda(0, \cdot) - v\|^2 + \left\|\sum_{ij} a^{ij}(t, \cdot)\frac{\partial v}{\partial y_i}\frac{\partial v}{\partial y_j}\right\| + \lambda^{2N}\|\nabla_y v\|^2 \leq 10,$$

the norms being taken in L^∞. That is, with

$$L(t) = -\sum_{ij} a^{ij}(t, y)\frac{\partial^2}{\partial y_i \partial y_j},$$

$(L(t))_{\lambda^{-2}}(u_\lambda(0, \cdot)) \leq 10$ for $|t| \leq \lambda$. Taking $t_\ell = \ell\lambda/N^2$, $\ell = 0, \ldots, N^2$, and applying Lemma 3 repeatedly, we obtain

$$\left(\sum_\ell L(t_\ell)\right)_{\lambda^{-2}}(u\ (0, \cdot)) \leq C_N,$$

which implies easily $(\overline{L}_\lambda)_{\lambda^{-2}}(u(0, \cdot)) \leq C_N'$. Thus, for a suitable C^1 function v we have

$$\lambda^{-2}|u_\lambda(0, y) - v(y)|^2 + \sum_{ij} a^{ij}(y)\frac{\partial v(y)}{\partial y_i}\frac{\partial v(y)}{\partial y_j} + \lambda^{2N}|\nabla v(y)|^2 \leq C$$

for all $y \in M$. In particular, $|v(0)| \leq C\lambda$. Moreover, let $\gamma : [0, \lambda] \to M$ be a Lipschitz path with $d\gamma/dt$ subunit for $\overline{L}_\lambda$. Our estimates on v and the definition of subunit vectors together yield

$$|v(\gamma(\lambda)) - v(\gamma(0))| \leq \int_0^\lambda |\langle dv, \dot\gamma(t)\rangle| dt,$$

while

$$\langle dv, \dot{\gamma}(t)\rangle^2 \le \sum_{ij} a^{ij}(y)\frac{\partial v}{\partial y_i}\frac{\partial v}{\partial y_j} + \lambda^{2N}|\nabla v(y)|^2 \text{ [with } y = \gamma(t)] \le C.$$

Therefore, $|v(\gamma(\lambda)) - v(\gamma(0))| \le C\lambda$, and it follows that $|v(y)| \le C\lambda$ for $y \in B_{\bar{L}_\lambda}(\lambda)$. The estimate defining v now shows that also $u_\lambda(0, y) \le C'\lambda$ for $y \in B_{\bar{L}_\lambda}(\lambda)$. Recalling that $|u_\lambda(t, y) - u_\lambda(0, y)| \le |t|$, we obtain $u_\lambda(t, y) \le C''\lambda$ for $(t, y) \in \{|t| \le \lambda\} \times B_{\bar{L}_\lambda}(\lambda)$. By definition of u_λ, this means that $\{|t| \le \lambda\} \times B_{\bar{L}_\lambda}(\lambda) \subseteq B_{L_\lambda}(C\lambda)$, as needed.

Sketch of the Proof of Lemma 3. We use induction on the dimension n. The result being trivial in $n = 0$, we assume Lemma 3 is known in R^{n-1} and deduce successively the following seven lemmas.

Lemma 3.1. Let $A = \sum_j a^j(x)\frac{\partial}{\partial x_j}$ *be a first-order operator in* R^{n-1}, *and let* $u, v, w : I \to C^1(R^{n-1})$, *where* I *is an interval of length* $\sim K^{-1/2}$. *Set* $u^0 = \frac{1}{|I|}\int_I u(t)dt$. *Then there exists* $u^+ \in C^1(R^{n-1})$ *for which*

$$K\|u^0 - u^+\|^2 + \|Au^+\|^2 + K^{-2N}\|\nabla_x u^+\|^2 \le C \sup_I\Big\{K\|u(t) - v(t)\|^2 + \left\|\frac{\partial v}{\partial t}\right\|^2 + K\|u(t) - w(t)\|^2 + \left\|\left(\frac{\partial}{\partial t} - A\right)w(t)\right\|^2 + K^{-2N}\|\nabla_x w\|^2\Big\},$$

the norms being taken in L^∞ (unit cube *in* R^{n-1}).

Lemma 3.2. Set $I_\ell = \{|t - \ell K^{-1/2}| \le 10K^{-1/2}\}$, I^+ = *union of consecutive* I_ℓ's,

$$L(t) = -\sum_{i,j=1}^{n-1} a^{ij}(t, x)\frac{\partial^2}{\partial x_i \partial x_j},$$

where a^{ij} *is smooth and* $(a^{ij}) \ge K^{-2N}(\delta_{ij})$. *Then given* C^1-*functions* $u(t, x)$ *on* $I^+ \times$ unit cube, u_ℓ *and* v_ℓ *on* $I_\ell \times$ unit cube, *we can find* $v \in C^1(I^+ \times$ cube) *so that*

$$K\|u - v\|^2 + \left\|\frac{\partial v}{\partial t}\right\|^2 + \left\|\sum_{ij} a^{ij}(t, x)\frac{\partial v}{\partial x_i}\frac{\partial v}{\partial x_j}\right\|$$
$$\leq C \max_\ell \left\{K\|u - v_\ell\|^2 + \left\|\frac{\partial v_\ell}{\partial t}\right\|^2 + K\|u - u_\ell\|^2 + \left\|\sum_{ij} a^{ij}(t, x)\frac{\partial u_\ell}{\partial x_i}\frac{\partial u_\ell}{\partial x_j}\right\|\right\}.$$

Here the norms on the left are taken in $L^\infty(I^+ \times \text{cube})$, *while those on the right are taken in* $L^\infty(I_\ell \times \text{cube})$.

Lemma 3.3. Set $A(t) = \sum_{j=1}^{n-1} a^j(t, x)\frac{\partial}{\partial x_j}$ *with* a^j *smooth, and let* I *be an interval of length* $\sim K^{-1/2}$. *Given* C^1*-functions* u, v, w *on* $I \times$ unit cube, *we can find* $u^\#$ *so that*

$$K\|u - u^\#\|^2 + \left\|\frac{\partial u^\#}{\partial t}\right\|^2 + \|A(t)u^\#\|^2 + K^{-2N}\|\nabla_x u^\#\|^2$$
$$\leq C\left\{K\|u - v\|^2 + \left\|\frac{\partial v}{\partial t}\right\|^2 + K\|u - w\|^2 + \left\|\left(\frac{\partial}{\partial t} + A(t)\right)w\right\|^2 + K^{-2N}\|\nabla_x w\|^2\right\},$$

the norms being taken in $L^\infty(I \times \text{cube})$.

Lemma 3.4. Let the unit cube in R^n *be partitioned into subcubes* $\{Q_\nu\}$ *of diameters* δ_ν, *where* $Q_\nu^* \cap Q_\mu^* \neq \emptyset$ *implies* $\delta_\nu \sim \delta_\mu$.

Let

$$L = -\sum_{ij} a^{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j}$$

for smooth functions a^{ij} *satisfying* $K^{-2N}(\delta_{ij}) \leq (a^{ij}(x)) \leq C\delta_\nu^2(\delta_{ij})$ *for* $x \in Q_\nu$. *Then given* $u \in L^\infty$ *and* C^1*-functions* v_μ *defined on* Q_ν^*, *we have*

$$L_K(u) \leq C \max_\mu \left\{K\|u - v_\mu\|_{L^\infty(Q_\mu^*)} + \left\|\sum_{ij} a^{ij}(x)\frac{\partial v_\mu}{\partial x_i}\frac{\partial v_\mu}{\partial x_j}\right\|_{L^\infty(Q_\mu^*)}\right\}.$$

Lemma 3.5. Assume $L = -\sum_{ij} a^{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \cdots$ *with* $\max_{ij}|a^{ij}| \geq C$, *and set*

$$L' = -\left(\frac{\partial}{\partial x_1}\right)^2 - K^{-2N}\Delta.$$

Then

$$c(L + L')_K(u) \le L_K(u) + L'_K(u).$$

Lemma 3.6. Set

$$L' = -\left(\frac{\partial}{\partial x_1}\right)^2 - K^{-2N}\Delta, \quad L = -\sum_{ij} a^{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j},$$

where $(a^{ij}(x))$ *is smooth and positive semidefinite. Then*

$$c(L + L')_K(u) \le L_K(u) + L'_K(u).$$

Lemma 3.7. Let

$$L = -\sum_{ij} a^{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j}, \quad L' = -\sum_{ij} a'^{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j},$$

and assume $\max_{ij}|a^{ij}| \ge C$. *Then*

$$c(L + L')_K(u) \le L_K(u) + L'_K(u).$$

For the most part, these lemmas are proved in strict analogy with Lemmas 1, 2, 3, 5, 7, 8, and 9 in the proof of the spectral decomposition theorem in [5]. Instead of L^2-norms, one uses L^∞-norms; and for

$$L = -\sum_{ij} a^{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j},$$

we work with $\left\| \sum_{ij} a^{ij}(x)\frac{\partial v}{\partial x_i}\frac{\partial v}{\partial x_j}\right\|_{L^\infty}$ in place of $\mathrm{Re}\langle Lv, v\rangle$.

In some ways, the present lemmas are easier than their L^2-analogues in [5]. For instance, where [5] uses the sharp Gårding inequality of [1], we use here merely the trivial estimate

$$\left\| \sum_{ij} b^{ij}(x)\frac{\partial v}{\partial x_i}\frac{\partial v}{\partial x_j}\right\|_{L^\infty} \le \left\| \sum_{ij} a^{ij}(x)\frac{\partial v}{\partial x_i}\frac{\partial v}{\partial x_j}\right\|_{L^\infty}$$

for $(b^{ij}) \le (a^{ij})$. Also, the Fourier integral operators of [5] are now replaced by simple changes of variable $y = \Phi(x)$.

It is now routine for a patient reader to reconstruct the proofs of Lemmas 3.2 through 3.7, given Lemma 3.1. Moreover, Lemma 3.7 implies Lemma 3 in R^n, just as Lemma 6.9 implies the spectral decomposition theorem in [5]. Consequently, the proof of Lemma 3 is reduced to Lemma 3.1. Since the L^2-analogue of 3.1 in [5] was proved using the spectral theorem, we have to make a new argument to show (3.1). Since space is limited, we omit all details of Lemmas 3.2–3.7 and Lemma 3.7 $\to$ Lemma 3 in R^n. We close our article with the proof of Lemma 3.1. First suppose $A = \partial/\partial x_1$, and let Ω denote the right-hand side of the estimate in Lemma 3.1.

Since $K\|u - v\|^2 + \left\|\frac{\partial v}{\partial t}\right\|^2 \le \Omega$, we have

$$K|u(t + h, x_1, x') - u(t, x_1, x')|^2 \le C\Omega$$

for $|h| \le K^{-1/2}$. Also $K\|u - w\|^2 + \left\|\left(\frac{\partial}{\partial t} - A\right)w\right\|^2 \le \Omega$, so that

$$K|u(t + h, x_1 + h, x') - u(t, x_1, x')|^2 \le C\Omega$$

for $|h| \le K^{-1/2}$; and $K\|u - w\|^2 + K^{-2N}\|\nabla_x w\|^2 \le \Omega$ so that

$$K|u(t, x_1, x') - u(t, x_1, y')|^2 \le C\Omega$$

for $|x' - y'| \le K^{-N}$. Now take

$$u^+(x_1, x') = \frac{1}{|I|}\int_I \int_{R^{n-1}} \psi(y_1, y')u(t, x_1 - y_1, x' - y')\,dt\,dy_1\,dy',$$

where $\psi(y_1, y')$ is a suitable approximate identity supported in $|y_1| \le K^{-1/2}$, $|y'| \le K^{-N}$. One checks easily using the previous estimates that

$$K\|u^0 - u^+\|^2 + \|Au^+\|^2 + K^{-2N}\|\nabla_x u^+\|^2 \le C\Omega,$$

as needed.

Next we pass to the general case. With $A = \sum_j a^j(x)\frac{\partial}{\partial x_j}$, we make a Calderón-Zygmund decomposition of the unit cube in R^{n-1}, stopping at Q_ν with diameter δ_ν if

(i) $\max_{x \in Q_\nu^*} \max_j |a^j(x)| \ge 10\delta_\nu$

or

(ii) $\delta_\nu \le CK^{-N}$.

In each Q_ν^* we can find a function u_ν^+ so that

$$K|u^0 - u^+|^2 + |Au^+|^2 + K^{-2N}|\nabla_x u_\nu^+|^2 \leq C\Omega \text{ on } Q_\nu^*.$$

For Q_ν arising from (i), this reduces by a change of variable to a slight variant of the case $A = \partial/\partial x_1$, in which the terms $K^{-2N}\|\nabla_x w\|^2$ and $K^{-2N}\|\nabla_x u^+\|^2$ are changed to $K^{-2N}\delta_\nu^{-2}\|\nabla_x w\|^2$ and $K^{-2N}\delta^{-2}\|\nabla_x u^+\|^2$— there is no trouble in adapting our discussion of $A = \partial/\partial x_1$ to this variant. If Q_ν arises from (ii), then we just take $u_\nu^+ = w$.

Now take a partition of unity $1 = \sum_\nu \phi_\nu$ with $\phi_\nu \in C_0^\infty(Q_\nu^*)$, $\|\nabla\phi_\nu\| \leq C\delta_\nu^{-1}$, and set $u^+ = \sum_\nu \phi_\nu u_\nu^+$. Clearly

$$K\|u^0 - u^+\|^2 \leq C\Omega.$$

We also have $Au^+ = \sum_\nu (Au_\nu^+)\phi_\nu + \sum_\nu (u_\nu^+ - u^0)(A\phi_\nu)$, while $|a^j(x)| \leq C\delta_\nu$ and $|\nabla\phi_\nu| \leq C\delta_\nu^{-1}$ in supp $\phi_\nu \subseteq Q_\nu^*$. Therefore,

$$\|Au^+\|^2 \leq C \max\{\|Au^+\|^2 + \|u_\nu^+ - u^0\|^2\} \leq C\Omega.$$

Similarly, $\nabla_x u^+ = \sum_\nu (\nabla_x u_\nu^+)\phi_\nu + \sum_\nu (u_\nu^+ - u^0)(\nabla\phi_\nu)$, while $|\nabla\phi_\nu| \leq \delta_\nu^{-1} \leq K^N$. So

$$K^{-2N}\|\nabla u^+\|^2 \leq C \max\{K^{-2N}\|\nabla_x u_\nu^+\|^2 + \|u_\nu^+ - u^0\|^2\} \leq C\Omega.$$

Thus

$$K\|u^+ - u^0\|^2 + \|Au^+\|^2 + K^{-2N}\|\nabla_x u^+\|^2 \leq C\Omega,$$

proving Lemma 3.1.

REFERENCES

1. Fefferman, C., and Phong, D. H. On positivity of pseudodifferential operators. *Proc. Nat. Acad. Sci.* 75 (1978):4673-4674.

2. Fefferman, C., and Phong, D. H. On the asymptotic eigenvalue distribution of a pseudodifferential operator. *Proc. Nat. Acad. Sci.* 77 (1980):5622-5625.

3. Fefferman, C., and Phong, D. H. On the lowest eigenvalue of a pseudodifferential operator. *Proc. Nat. Acad. Sci.* 76 (1979): 6055-6056.

4. Fefferman, C., and Phong, D. H. Symplectic geometry and positivity of pseudodifferential operators. *Proc. Nat. Acad. Sci.* 79 (1982):710-713.

5. Fefferman, C., and Phong, D. H. The uncertainty principle and sharp Gårding inequalities. *Comm. Pure Appl. Math.* 34 (1981): 285-331.

6. Hörmander, L. Hypoelliptic second-order differential equations. *Acta Math.* 119 (1967):147-171.

7. Kohn, J. J. Pseudodifferential operators and hypoellipticity. *Proc. Symp. Pure Math. AMS* 23 (1973):61-69.

8. Menikoff, A., and Sjöstrand, J. Eigenvalues for hypoelliptic operators and related methods. *Proc. Int. Cong. Math. Helsinki* (1978):797-801.

9. Olenik, O., and Radkevitch, E. Second-order equations with non-negative characteristic form.

10. Rothschild, L. and Stein, E. M. Hypoelliptic differential operators and nilpotent groups. *Acta Math.* 137 (1976):247-320.